Catalytic substitution
reaction breaks
the rules of organic
chemical synthesis
PAGES 438 & 447

MOLECULAR CONTROL

Challenge anti-Semitism

A wave of anti-Jewish prejudice is once again washing over schools and universities. There is no excuse not to call out this vile behaviour.

A dark shadow is stalking the land. Anti-Semitism is once again showing itself and must be confronted head-on and stamped out wherever it is found. Discrimination against any individuals and groups goes against the values of science, human decency and this journal. Sunlight remains the best disinfectant, and one of the places where the infection of anti-Semitism still thrives is on some university campuses. If researchers who work there are not aware of this and angry about it, then they have not been paying sufficient attention.

There are many examples and many responses. Just last week, South Carolina took another step towards a law that would make it the first US state to set a legal definition of anti-Semitism. Supporters of the move argue that it is needed to help university administrators to combat a rising tide of hate against Jewish staff and students. Critics say it impinges on free speech. That this is being discussed at all in 2018 speaks volumes.

One does not have to delve too deeply into modern international politics to find clear examples of a renewed and ugly mood of hostility towards Jews. The prime minister of Hungary, Viktor Orbán, secured a fourth term in office earlier this month on an aggressively nationalist ticket widely criticized for anti-Semitic images and messages. And last week saw the grim spectacle of British MPs standing in Parliament and reading out some of the intolerable anti-Semitic abuse they have received as part of a highly unusual debate on the subject.

It would be surprising if this political climate did not embolden anti-Semites on campus. Reliable numbers are hard to come by. In a report last year, the European Union Agency for Fundamental Rights pointed to gaps in data about anti-Semitic incidents across the continent, which it intends to address with a survey later this year. But some estimates do indicate that there has been a surge in places, including schools and universities.

In the United States, the Anti-Defamation League reported in February, such incidents on university campuses increased by 89% in 2017, to 204. Surveys in the United States and the United Kingdom highlight that many Jewish students find the atmosphere on campus intimidating. This is hardly surprising, given that one of the most common offences is to draw a swastika on a wall.

Anti-Semitism — prejudice and violence against individuals and communities — is distinct from legitimate criticism of Israeli policy. It is perfectly possible to argue the rights and wrongs of international politics without hate speech.

The following is a widely accepted definition from the International Holocaust Remembrance Alliance: "Anti-Semitism is a certain perception of Jews, which may be expressed as hatred toward Jews. Rhetorical and physical manifestations of anti-Semitism are directed toward Jewish or non-Jewish individuals and/or their property, toward Jewish community institutions and religious facilities."

According to this definition, examples of anti-Semitic behaviour include (but are not limited to) furthering the myth that Jews are engaged in a shadowy conspiracy to control events, and holding Jews collectively responsible for the actions of the government of Israel. It should not need saying that this is as preposterous as holding Muslims collectively responsible for the actions of the Syrian government.

Nature has a long history of highlighting and confronting anti-Semitism. In an editorial in December 1935, this journal warned that an "anti-Semitic clique" was trying to seize control of the Germany-based Astronomical Society (*Nature* **136**, 927–928; 1935). Less than

> *"Many Jewish students find the atmosphere on campus intimidating."*

a month later, we noted that discrimination against Jews and other "non-Aryans" meant that "Germany stands condemned as guilty of a persecution no less barbarous and an intolerance as rigid and as crass as any that figure in the annals of the Middle Ages" (*Nature* **137**, 16; 1936).

It is sad and worrying that we feel the need to highlight the point again. But we are confident we can rely on *Nature*'s readers to challenge anti-Semitism whenever and wherever it occurs — in their universities, on campus, at social occasions, or on the street — just as we can rely on readers, as we have frequently urged them, to challenge those who express their hatred of people of colour, women, Muslims, immigrants, the gay and transgender community and many others. ∎

A sting in the tale

As climate officials meet to swap stories, only emissions cuts will guarantee a happy ending.

Countries will initiate the first formal review of progress under the 2015 Paris climate pact at the United Nations climate talks in Bonn, Germany, next week. According to the UN, the 'Talanoa Dialogue' aims to "share stories, build empathy and to make wise decisions for the collective good. The process of Talanoa involves the sharing of ideas, skills and experience through storytelling."

Delegates will no doubt come to the table with countless tales of deployment of clean-energy technologies and initiatives to help communities to prepare for a warming world. But the process must prompt serious self-reflection from policymakers on how far they are falling short of their 2015 climate commitments — and what it will take for them to put the world on a track to real sustainability.

As we discuss in a News Feature this week (page 422), there is encouraging news. The nascent clean-energy industry has found its feet. Wind and solar sources are already used ahead of fossil fuels in many places, and renewables will become only more competitive as technology prices fall in the years to come. Indeed, at least one

energy consultancy predicts that installing solar panels in 2030 will be a cheaper way of generating energy in some places than continuing to shovel coal into the furnaces of ageing power plants. Similarly, electric vehicles are likely to become cheaper to buy and run than petrol-powered incumbents over the next decade.

These genuine tipping points could disrupt energy markets and potentially enable much faster change than anybody anticipates today. From this perspective, it is surely only a matter of time before clean-energy technologies come to dominate. But time is precisely what humanity does not have — not if it wants to meet the Paris goal of limiting warming to 1.5–2 °C above pre-industrial levels.

Renewables now capture the bulk of global investments in new power installations each year, covering much of the annual growth in energy demand. That leaves a lot of existing infrastructure that continues to burn oil, coal and gas — with the financial support of governments. Despite efforts to reduce fossil-fuel subsidies, a recent study found that governments still handed over some US$330 billion of subsidies in 2015 (J. Jewell *et al. Nature* **554,** 229–233; 2018).

The consequences are clear. With nearly 1 °C of warming in the bag, the world is already experiencing unwanted effects, such as extreme weather events. These will continue to mount unless and until humanity slashes its greenhouse-gas emissions. A sober look at the numbers suggests that this task will be difficult — if not impossible — without radical interventions to deliberately steer energy producers and users towards sustainable options.

That takes political commitment, which will be one focus of next week's dialogue. The news on this score is mixed at best. US President Donald Trump is promoting a retrograde energy agenda, and has vowed to pull the United States out of the Paris agreement. (Still, despite their rhetoric and ambitions, Trump and his allies have yet to halt coal's probably inevitable decline in the face of cheap natural gas and renewables.)

Support for the Paris agreement remains high outside the United States, but has its limits. Estimates suggest that actions on emissions so far have probably shaved off 1 °C or so from the projected warming this century, but the world remains on course for a rise of well over 3 °C (see page 424). That's true even if countries fulfil their current emissions pledges, which isn't likely.

Perhaps the best news is that developing countries — including China and India, plagued by air pollution in many urban areas — have come to view clean energy through the lens of public health and air quality. Policies on climate and sustainable development that promote clean, low-carbon energy go hand in hand.

"Support for the Paris agreement remains high outside the United States."

Another way to measure progress is to look at where the money is going. Given the scale of the challenge ahead, the goal of policymakers must be to align investments across the climate landscape, from energy efficiency and carbon-free energy sources to green buildings, cities and other infrastructure.

Ultimately, human influence on the climate comes down to one thing: what volume of carbon dioxide and other greenhouse gases is pumped into the atmosphere. This means that ramping up renewables to meet energy demand simply isn't enough. Reducing carbon emissions means making painful choices: halting new investments in the exploration and production of fossil fuels, and then closing down existing facilities. It won't be easy, but eventually that is a story that must be told. ∎

Visual science

Results from a Nature *photo competition show the power of images.*

If writing about music is like dancing about architecture, then writing about photography seems a tough ask indeed. Luckily, not everyone sees it that way — and certainly not when it comes to science. As the celebrated US photographer Berenice Abbott wrote in 1939: "To obtain wide popular support for science, to that end that we may explore this vast subject even further and bring as yet unexplored areas under control, there needs to be a friendly interpreter between science and the layman. I believe that photography can be this spokesman, as no other form of expression can be."

In a Careers article this week (page 525), we celebrate both the spirit and the letter of Abbott's words. During March, we asked readers to send in their own photographs to our 2018 #ScientistAtWork contest. Some 330 readers did so; entries ranged from images of researchers at work around the world to depictions of their actual work, and were assessed by a panel of *Nature* journalists and art editors.

The overall winner was a striking and beautifully framed shot of marine biologist Callie Veelenturf kneeling beside a leatherback sea turtle in Equatorial Guinea, taken by her colleague Jonah Reenders. The picture is worth a thousand words to us — or more — and a year's personal subscription to *Nature* for Reenders and Veelenturf.

Overall, the competition entries show the sheer diversity of modern research perhaps better than any words can. If teachers or scientists who visit schools want to demonstrate that the job of a scientist is varied and sometimes extraordinary, then this collection is a good start.

Photography and research have a history that started long before Abbott turned her lens to science. The technique started as a scientific endeavour — indeed, some artists rejected the medium because it looked too lifelike, in a way that a painting or an engraving never could. Still images have recorded, shared and prompted discovery ever since. From the very small, taken through microscopes, to the very far away, imaged by telescopes, photographs comprise an essential and valued contribution to the scientific record.

Can the camera lie? It can certainly bend scientific truth. James Nasmyth, the British engineer and amateur astronomer, produced what looked like impressive photographs of the Moon back in 1874. In fact, the photos were of detailed plaster models of the lunar surface that Nasmyth built from careful telescope observations. It took almost another century before genuine photos of the Moon emerged, under just as bizarre circumstances — when British astronomers used a giant radio telescope and a borrowed fax machine to hack into the 1966 signals sent from the Soviet space probe Luna 9.

Two more photos sent from space proved hugely influential. On Christmas Eve 1968, the *Apollo 8* astronaut William Anders was stunned by what he saw from his window as they flew around the Moon: "Oh my God! Look at that picture over there! Here's the Earth coming up. Wow, is that pretty … You got a color film, Jim?" The resulting image of our living planet suspended in space above the barren rocky surface of the Moon, called *Earthrise*, was described by the wildlife photographer Galen Rowell as "the most influential environmental photograph ever taken".

More than two decades later, another picture of Earth — this time taken by the speeding Voyager 1 probe from 6 billion kilometres away — showed our planet as a mere speck of light amid vast surroundings and was named the *Pale Blue Dot*. (Technically, the picture is a composite made up of three separate frames, each taken through a different colour filter.)

Neutrons, viruses and the (former) planet Pluto: all were found thanks to photographs. But there is something else that runs through the images sent to *Nature*, and not something that is easily captured in academic prose. Spirit, perhaps, or joy; maybe what Abbott called "vivification of the visual image, the warm human quality of imagination added". Do take a look and decide for yourselves. ∎

A lab co-op helps young faculty members to thrive

Linking a lab with others fosters crucial camaraderie, collaboration and productivity, writes **Rebecca Heald**.

When I started my lab at the University of California, Berkeley, two decades ago, what terrified me most was the thought that I alone was responsible for everything — from formulating successful PhD-thesis projects to picking the right freezer. Fortunately, my stress was mitigated because, within a year, two new assistant professors, Matt Welch and Karsten Weis, were hired and given lab space next to mine. Although we all focused on different areas of cell biology, we shared common interests and values and quickly saw benefits in joining forces. We called our joint groups the Trilab.

Matt, Karsten and I saved space by sharing chemical and microscopy rooms, and saved money by pooling equipment. We even tore down a wall to create a joint lunch room, where members of our groups could socialize and discuss projects. We held joint lab meetings weekly, which provided a greater sense of progress and exposed trainees to a wider range of topics, techniques and expertise than any of our labs could have done on their own. The Trilab arrangement also meant that, as lab heads, we had close colleagues with whom to bounce around ideas, trade grant proposals, provide encouragement and have some fun. (Our annual, department-wide Halloween party has become famous: every year, our lab members squeeze enough limes to make 20 litres of margaritas.) Without any kind of master plan, a nurturing scientific and social environment emerged.

What happened for me serendipitously is something that I urge other faculty members, particularly younger ones, to seek out deliberately.

A tight network of colleagues facilitates collaboration; with that come the courage and capacity to tackle interdisciplinary projects. Early on, Karsten and I realized that our labs had complementary knowledge and the approaches necessary to make real headway in an area dominated by established groups. Thanks to funding from the US National Institutes of Health, and the hard work of many students and postdocs over the 12 years that followed, we were co-corresponding authors on 9 research articles applying new tools we had developed for microscopy and chemical biology. Among other creative approaches, we found a way to mimic chromosomes by coupling a single protein to porous glass beads. This was sufficient to cause the chromosome-transport apparatus — the spindle — to assemble in egg-cell extracts, which revealed surprising mechanisms about how cell division is orchestrated. Our collaboration bolstered both of our tenure cases, and the department celebrated our success without attempting to parse out who deserved what portion of credit.

Emotional support was just as important. Matt and I commiserated in each other's offices over failed funding applications more times than I want to remember. We were each willing to do our share or more to make sure all three labs succeeded. Matt and Karsten, in one instance, applied for a grant to buy a microscope that would benefit all of us.

The biggest challenge is finding the right colleagues with whom to form the group. Proximity is key. I advise postdocs who want to follow conventional paths in academia to prioritize jobs in departments that are hiring lots of junior faculty members, and to avoid institutions — even prestigious ones — that force assistant professors to compete with each other. That makes everyone in the lab miserable.

Matt, Karsten and I had each previously experienced collaborative environments — Matt at the University of California, San Francisco; me at the European Molecular Biology Laboratory in Heidelberg, Germany; and Karsten at both — which helped us conceive of the Trilab. We were exceedingly fortunate to be at the same career stage at the same place and time. But there are other elements necessary for building a productive lab network. These start in the individual lab.

Even when forming a co-op is impractical, the mindset behind it improves the health of individual labs. A principal investigator should set the right tone by designing projects to be complementary rather than overlapping or competing, so that lab members take ownership of them and the environment is supportive, not antagonistic. This might slow progress, but it increases motivation. Leaders should also take care to capitalize on the range of experience in the group, and treat lab members with equal respect, no matter what their background or career aspirations. Scientists who help each other to collect, analyse and quantify data increase the rigour of the lab's work as a whole.

THE
MINDSET
BEHIND A CO-OP
IMPROVES
THE HEALTH
OF INDIVIDUAL
LABS.

Moreover, generosity is contagious. In a large group, not everyone is going to buy in, but good will can become the normal state. Twenty years on, Karsten's lab has moved to Zurich, Switzerland, but the Trilab has become the Tetralab, with two other groups joining me and Matt.

There are ways of networking even when labs are farther apart. One colleague, biophysicist Eva Nogales, started a monthly, ongoing junior-faculty lunch club that spanned biology, chemistry and physics. We have also benefited from mini-retreats; about twice a year, we invite nearby labs with overlapping interests for a snack-filled afternoon of short talks and brainstorming.

Celebrating together is key to making a network strong. In our group, we reward ourselves for a manuscript submission. In my lab, at least, it is a huge, multi-year accomplishment to finish a paper, which more often than not is initially rejected. The long and arduous process of publishing a paper has formally begun! Group support is essential in both good and bad times.

Once the community is set up, it self-propagates. What benefits the larger group also benefits the individual lab, and vice versa. A network of human interactions is central to progress and success. Every researcher should make establishing these support systems a priority. ∎

Rebecca Heald *is a professor of cell and developmental biology at the University of California, Berkeley.*
e-mail: bheald@berkeley.edu

SEVEN DAYS

The news in brief

Dengue vaccine

The sole approved vaccine for dengue fever must be given only to people who have had a past infection of the disease confirmed by a laboratory test, a World Health Organization advisory group recommended on 19 April. The move follows the finding that although Dengvaxia, manufactured by drug firm Sanofi, is effective in populations with high levels of dengue when viewed as a whole, it can cause severe dengue fever in individuals who have never been infected by the mosquito-borne virus. Lab testing for past infections could complicate, if not stall, routine vaccination efforts because the available tests for dengue are poor. The advisory group says that developing better tests should be a priority. It estimates that there are 50 million to 100 million cases of dengue per year, mainly in Asia.

PAUL HENNESSY/ALAMY

Planet hunter sets off for distant worlds

NASA's Transiting Exoplanet Survey Satellite (TESS) launched (pictured) from Cape Canaveral, Florida, on 18 April on a two-year mission to discover planets beyond the Solar System. TESS is meant to build on the work of the Kepler spacecraft — which has found thousands of exoplanets since 2009 — by hunting worlds that lie within about 90 parsecs (300 light years) of Earth. The planets TESS discovers will form the basis for decades of astronomical study.

PUBLISHING

Copyright deal

Academic publishers Springer Nature, Cambridge University Press and Thieme have reached an agreement with ResearchGate, the world's largest scholarly social network, on the responsible sharing of journal articles. (*Nature*'s news team is editorially independent of its publisher, Springer Nature.) The agreement, announced on 19 April, allows the three publishers to track and record user-uploaded content on the platform. ResearchGate will also ensure that when it is alerted to articles that infringe copyright rules, the content is promptly removed. In recent months, the site has disabled access to more than 1 million papers in response to take-down requests from publishers. But many worry that large numbers of copyright-infringing articles are still freely accessible on the platform. Publishers Elsevier and the American Chemical Society have taken the Berlin-based company to court to clarify copyright responsibilities. The first hearing was held on 18 April in Munich, Germany.

Regional research

A citation database that promises to cover many more regional journals than other existing databases do launched on 24 April. Called 1findr and run by information-systems company 1Science in Montreal, Canada, the service is free for users, but a subscription version will offer extra features, such as allowing users to download up to 50 articles with a single click. The database so far contains records of 90 million peer-reviewed journal articles. The platform's creators aim to eventually index all of the world's peer-reviewed papers — regardless of where and in what language they are published. The database is aimed mostly at librarians, academic institutions, and researchers conducting data-mining projects.

POLICY

Therapy-ban review

Singapore's bioethics advisory committee announced on 19 April that it is reviewing its research ban on mitochondrial replacement — a controversial technique that creates embryos using DNA from three people. The method seeks to allow women who carry disease-causing mutations in their mitochondrial DNA to have healthy babies. It works by moving nuclear genetic material from an egg with faulty mitochondria into a healthy donor egg that has had its nuclear DNA removed. Although such research is currently banned in Singapore, it has been conducted in Mexico, the Ukraine and China, and is approved in the United Kingdom. The Singaporean committee is inviting public feedback until 15 June.

UK genomic report

The UK House of Commons Science and Technology Committee tackled genomics and gene editing in a report released on 20 April. The report notes that budget cuts to the country's health-care system could endanger efforts to incorporate genomics into patient care and recommends raising public awareness of genomic medicine. The committee also argues that government-collected genomic data should be better integrated

with other health-care data held by the public and private sectors. The government should also monitor genome-editing efforts for obstacles to the technology's development, it says.

Epilepsy drug

A drug containing marijuana components has, for the first time, been recommended for approval by an advisory panel to the US Food and Drug Administration (FDA). On 19 April, the panel voted unanimously to approve Epidiolex, an anti-epilepsy drug made from cannabidiol, which is a chemical in marijuana that does not cause psychoactive effects. The drug is made by GW Pharmaceuticals in Cambridge, UK, and it could be approved to treat two epilepsy disorders, Lennox–Gastaut syndrome and Dravet syndrome. In clinical trials, Epidiolex reduced the number of seizures in some people by more than 50%. However, the panel's recommendations are not binding, and the FDA is expected to make a final decision by the end of June.

New NASA chief

NASA's new leader is James Bridenstine, a Republican member of Congress from Oklahoma with a strong interest in commercial spaceflight. The US Senate confirmed him on 19 April, almost 8 months after US President Donald Trump nominated him for the job, and 15 months after the departure of NASA's previous chief, former astronaut Charles Bolden. Bridenstine's confirmation was held up partly by political infighting, but also over questions about whether he supports climate-change research. Bridenstine (**pictured**) takes over an agency newly tasked with returning astronauts to the Moon but with few financial resources towards that goal.

Salk investigation

On 21 April, the Salk Institute for Biological Sciences in La Jolla, California, put cancer researcher Inder Verma on administrative leave while it investigates unspecified allegations against him.

The institute has enlisted an outside party to look into the claims, some of which were raised by a reporter from *Science*, according to a statement by Dan Lewis, chair of the institute's board of trustees. Through his lawyer, Verma has denied acting inappropriately. "I have never used my position at the Salk Institute to take advantage of others," he said in a statement to *Nature*. "I have also never engaged in any sort of intimate relationship with anyone affiliated with the Salk Institute. I have never inappropriately touched, nor have I made any sexually charged comments, to anyone affiliated with the Salk Institute. I have never allowed any offensive or sexually charged conversations, jokes, material, etc. to occur at the Salk Institute."

Sentenced scientist

Ahmadreza Djalali, an emergency-medicine researcher who was sentenced to death in Iran in October last year for spying, has sent an open letter to the scientists who have supported him. Djalali was working at the Karolinska Institute in Stockholm when he was detained during a trip to Iran on 26 April 2016. He was convicted of espionage following a trial in Iran's revolutionary court. Djalali says he was asked to spy by both Iranian and European intelligence services, but he maintains that he never accepted. In the letter, which marks the two-year anniversary of his arrest, he thanks the scholars who helped him but says that he is having serious health problems in prison. Djalali has appealed against his sentence at Iran's supreme court, and Nobel laureates sent a letter to the Iranian government about the case last November; neither action has yet received a response.

Biomedicine boost

Singapore will increase spending on health and biomedical research, it was announced on 17 April. The country's National Medical Research Council awarded 60 million Singapore dollars (US$45.6 million) over five years to projects that will study three diseases common in Singapore and southeast Asia: lung cancer, virus-induced cancers and age-related macular degeneration. Another scheme will provide clinicians with up to 100,000 Singapore dollars in seed funding to develop new medical technologies and devices. The latest measures support research translation, a key focus of Singapore's five-year Research, Innovation and Enterprise 2020 Plan.

Methane on Mars

A Russian–European Mars probe that is poised to solve one of the most controversial mysteries in Martian science began making scientific observations on 21 April. Researchers hope that the Trace Gas Orbiter (TGO) will resolve a debate over the origins of methane on the red planet, which could be a signature of life. See page 419 for more.

↻ **NATURE.COM**
For daily news updates see:
www.nature.com/news

TREND WATCH

The number of adults living with two or more chronic mental or physical conditions is increasing worldwide. An analysis of existing research compiled by the UK Academy of Medical Sciences found that 'multimorbidity' is most common in older people, but evidence on the extent and causes of the problem is fragmented. For instance, some studies reveal a clear association between multimorbidity and poverty. But some research suggests that certain chronic conditions are more common in wealthy people.

CHRONIC-ILLNESS BURDEN

Age is strongly associated with multimorbidity — having two or more chronic illnesses — but economic status is also a factor.
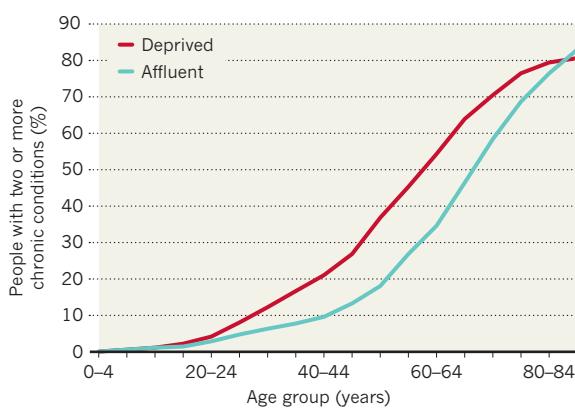
NEWS IN FOCUS
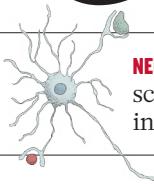
Flooding in Bangladesh could become more common as global temperatures rise.

CLIMATE

Telltale warming likely to hit poorer countries first

Climate–inequality tool reveals how quickly abnormal weather will appear around globe.

BY QUIRIN SCHIERMEIER

Nations such as Bangladesh and Egypt have long known that they will suffer more from climate change than will richer countries, but now researchers have devised a stark way to quantify the inequalities of future threats.

A map of "equivalent impacts", revealed at the annual meeting of the European Geosciences Union this month in Vienna, shows that global temperatures would have to rise by a whopping 3 °C before most people in wealthy nations would feel departures from familiar climate conditions equal to the changes that residents of poorer nations will experience under moderate warming.

The Paris climate agreement, adopted by 195 countries in 2015, aims to limit the rise in global mean temperature to 1.5–2 °C above pre-industrial levels. The world has already warmed by one degree or so — and, since 1900, the mean number of record-dry and record-wet months each year has also increased.

But the effects of global warming are uneven, and poor regions in the tropics and subtropics are thought to be most vulnerable, for several reasons. They have limited financial resources with which to prepare for shifts in temperature and precipitation, and they are expected to face bigger changes in climate than will countries in the mid-latitudes. Researchers have had difficulty quantifying those inequalities because the impacts of climate change depend on many factors, such as future economic growth ▶

and technological progress, which are hard to forecast.

Luke Harrington, a climate researcher at the University of Oxford, UK, took a different approach by developing the concept of equivalent impacts, which doesn't specify societal consequences. Instead, it focuses on quantifying the uneven distribution of extreme weather around the globe.

Harrington looked at changing patterns of extreme daily heat and rainfall in global climate projections based on fast-rising greenhouse-gas emissions. He then determined how much warming was required for a clear climate-change signal — such as extreme temperatures or precipitation — to emerge from the 'noise' of natural climate variability at each spot on the globe. The resulting maps show how quickly regional changes in weather extremes will manifest in response to different levels of global warming.

"I wanted to wrap numbers around the

"I wanted to wrap numbers around the unevenness of impacts."

unevenness of impacts," he says. "Climate-mitigation policies focus on a global threshold — but global mean temperature isn't a very meaningful metric to assess what climate change might mean in specific parts of the world," says Harrington, whose work has not yet been accepted for publication.

For changes in regional heat extremes, the pattern is particularly stark. Africa, large parts of India and most of South America are likely to experience changes clearly attributable to climate warming early on, after a 1.5 °C increase in global temperatures. But mid-latitude regions — where most greenhouse gases are produced — won't see such pronounced changes until global temperatures rise by 3 °C or so.

"This is an elegant way to tie global climate

targets and regional impacts," says Erich Fischer, a climate scientist at the Swiss Federal Institute of Technology Zurich, who was not involved in the study. He says that the model would need to be adapted to include metrics of specific climate-change impacts, such as those on human health and food security, for it to be useful for planning adaptation efforts or for informing international climate-finance programmes. Some proposed schemes would compensate poor countries for climate-change-related harm.

The equivalent-impacts index, says Fischer, could help quantify how climate change will affect different countries, because it focuses on identifying when they will start to face weather outside their natural variability.

"Our study provides a framework," say Harrington. "We want to know what information others care about most, then we can start to look at metrics of more-specific climate impacts." ■

Early success fuels further grants

Researchers who just miss cut-off for postdoc grant fall behind those who narrowly qualify.

BY HOLLY ELSE

The career-defining effect of winning a postdoctoral research grant has been laid bare in an analysis of thousands of young researchers' professional trajectories. The work compared the fate of junior scientists in the Netherlands who just met the bar to qualify for post-PhD research funding with that of people who just missed out on the money. The successful group went on to secure more than twice as much research funding in the subsequent eight years, the analysis found. And the grant-winners were also 50% more likely to become professors than were the ones who fell short. The study was published on 23 April (T. Bol *et al. Proc. Natl Acad. Sci USA* https://doi.org/cnrr; 2018).

What is most striking is that winning the initial grant did not have any effect on the scientists' publications or academic impact

THIN LINE
Researchers who just qualified to win a certain early-career grant went on to receive much more research funding in the years afterwards than did those who just missed out, an analysis finds.

NWO, Netherlands Organization of Scientific Research; ERC, European Research Council

in the following years, says Shulamit Kahn, an economist at Boston University in Massachusetts. Funders often consider previous

awards when making decisions about whom to give money to. "Why are they doing this if it doesn't increase productivity?" asks Kahn,

---

MORE ONLINE

NATURE BRIEFING
Save time — get the daily *Nature* newsletter free in your inbox go.nature.com/savetime

MORE NEWS
● Grassland plants show surprising appetite for carbon dioxide go.nature.com/2qxgfs3
● Microscope spies on living cells inside the body go.nature.com/2hnjo12
● 'Warm transplants' save livers and lives go.nature.com/2hopveh

NATURE PODCAST
The ethics of model minds, and an updated view of a chromosome-protecting enzyme nature.com/nature/podcast

adding that every funding body should be looking at the effect of their grants.

"If scientists are dissuaded from science by lack of funding, then the investment in scientific training becomes a sunk cost," says economist Donna Ginther of the University of Kansas in Lawrence.

COMMON TREND

Previous studies have made similar findings about the effects of early-career grants on later success, but the authors of the latest work say that they compared the fate of researchers with similar abilities in a way that no one else has. Earlier this month, Ginther published the results of a similar analysis, which found that securing a specific early-career fellowship from the US National Institutes of Health increases a researcher's chance of winning more grants from the funder (M. L. Heggeness *et al.* NBER Working Paper No. 24508; National Bureau of Economic Research, 2018).

The Dutch study, led by sociologist Thijs Bol at the University of Amsterdam, draws on data from the Netherlands Organization of Scientific Research (NWO), the country's national research council. The NWO operates a three-stream funding scheme that sets aside a total of €150 million (US$183 million) a year for scientists in the early, middle and established stages of their careers. Bol and his colleagues tracked more than 4,000 researchers who applied for the scheme's early-career grant between 2002 and 2008.

They looked at the grant-application scores of those academics, and tracked whether they went on to secure a mid-career grant from the funder in the following eight years. They also counted any grants from the European Research Council won between 2005 and 2016.

For around 1,400 of the early-career applicants, the researchers sourced data from article database Scopus about their publication and citation records before, during and after the time period of the NWO grants. They also determined how many of them had become full professors by 2018.

They found that candidates slightly above and below the funding cut-off had different career trajectories, even though their publication and citation records remained similar.

Researchers who ranked just above the threshold secured €180,000 in research funding over the next 8 years — more than twice as much as those just below it (see 'Thin line'). This was partly because researchers who lost out on the initial grant were less likely to apply for future funding.

"There is a group of very young talented scholars who have bad luck," says Bol. "They do not get the same resources to bring their ideas to life." ∎
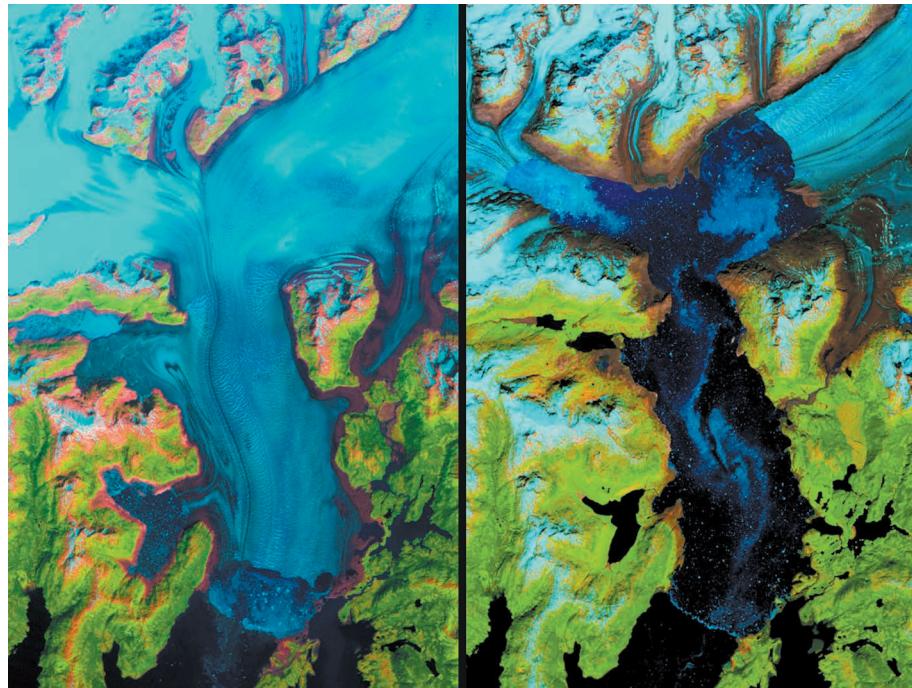
The ongoing melting of Alaska's Columbia glacier is shown in these Landsat images from 1986 and 2017.

EARTH OBSERVATIONS

US government reviews data fees

Images from Landsat satellites and agricultural–survey programme are freely available to scientists — for now.

BY GABRIEL POPKIN

The US government is considering whether to charge for access to two widely used sources of remote-sensing imagery: the iconic Landsat satellites operated by the US Geological Survey (USGS) and an aerial-survey programme run by the Department of Agriculture (USDA).

Officials at the Department of the Interior, which oversees the USGS, have asked a federal advisory committee to explore how putting a price on Landsat data might affect scientists and other users; the panel's analysis is due later this year. And the USDA is contemplating a plan to institute fees for its data as early as 2019.

Researchers who work with the data sets fear that changes in access could impair a wide range of research on the environment, conservation, agriculture and public health. "It would be just a huge setback," says Thomas Loveland, a remote-sensing scientist who recently retired from the USGS in Sioux Falls, South Dakota.

The Landsat programme began with one

satellite in 1972, and has launched another seven since. Together, they have produced the world's longest-running data set of satellite images. The two current probes take pictures at a resolution of 30 metres up to every 8 days.

Until 2008, researchers had to buy Landsat images — and they often designed studies to limit data costs, Loveland says. "You would buy as few images as you possibly could to get an answer."

Since the USGS made the data freely available, the rate at which users download it has jumped 100-fold. The images have enabled groundbreaking studies of changes in forests, surface water and cities, among other topics. Searching Google Scholar for "Landsat" turns up nearly 100,000 papers published since 2008.

A USGS survey of Landsat users released in 2013 found that the free distribution of imagery generates more than US$2 billion of economic benefit annually — dwarfing the programme's current annual budget of roughly $80 million. More than half of the nearly 13,500 survey respondents were academics, and the majority lived outside the United States.

In July 2017, officials at the Department of the Interior asked a committee of external advisers to study whether Landsat's costs could be recovered from users. The panel is preparing a white paper for release this year. "It's a serious discussion," says committee member Rebecca Moore, director of engineering at Google's Earth Engine.

Loveland says that trying to make the Landsat programme pay for itself could backfire: charging for data would probably lower usage, and the administrative costs of handling payments would eat into any revenue. "It costs a lot of money to charge money," he says.

The last time the federal advisory committee examined whether to reinstate fees for Landsat data, in 2012, it concluded that "Landsat benefits far outweigh the cost". Charging money for the satellite data would waste money, stifle science and innovation, and hamper the government's ability to monitor national security, the panel added.

Then there is the USDA's National Agricultural Imagery Program. Since 2003, it has hired companies to gather images of Earth's surface using aircraft, covering the entire United States at least once every three years. The resulting pictures have a resolution of 1 metre, enabling scientists to detect individual trees and buildings.

The data are "a critical component to [land] management here in the West", says April Hulet, an ecologist at the University of Idaho in Moscow who uses the images to study invasive plant species and fire risk. If the USDA began charging for the information, Hulet says, she would probably pay — if she could afford it.

The USDA is considering whether to license the data for a fee starting in 2019, according to minutes from a November 2017 meeting of an interagency panel that oversees US geospatial policy. The USDA hopes to have a draft plan ready by the end of summer, and then post it for public comment, says Denny Skiles, director of the department's Aerial Photography Field Office in Salt Lake City, Utah.

There are no perfect substitutes for images from Landsat or the USDA programme. Companies such as Planet and DigitalGlobe collect high-resolution satellite images and give scientists free access to some of those data. But buying commercial imagery that covers large areas or long periods is too expensive for many researchers. And although the European Space Agency's Sentinel-2 satellites provide free global imagery at resolutions up to 10 metres, they cannot match Landsat's 46-year record, says Martin Herold, a remote-sensing expert at Wageningen University in the Netherlands.

"The longer and more dense the archive," he says, "the more valuable it becomes." ∎

POLICY

Rent rise frustrates EU drug agency

European Medicines Agency hits relocation stumbling block.

BY INGA VESPER

Amsterdam's red-hot property market might threaten the mission of Europe's drug regulator. With less than a year to go before the European Medicines Agency (EMA) must leave its London headquarters because of Brexit, the agency is facing an unexpected rent increase that could cut into its budget for approving new medicines and overseeing clinical trials.

The multimillion-euro rent increase, which is spread over two decades, was revealed — along with other problems with the agency's move to Amsterdam — in EMA board-meeting minutes that were released this month.

The developments threaten to drain crucial resources, financial and otherwise, at a time when the agency's operational capacity is uncertain owing to expected staff losses caused by the move, say agency officials and observers.

The EMA is engaging in frantic business planning to ensure the move takes place "with as little interruption as possible", says deputy executive director Noël Wathion.

But the agency might soon cut some crucial functions for science, including advising on early-stage drug research and implementing a directive to ensure good practice in clinical trials, according to contingency plans drafted by the agency to prepare for the move.

"We had a very open and frank discussion with the Dutch colleagues and we reached an agreement."

The EMA, whose roughly 900 staff members oversee the safety of European medicines and offer scientific advice, is scheduled to move from London to Amsterdam in March 2019. Its new location was chosen last November by the European Union, from bids submitted by several countries, including Italy, Sweden and Slovakia. The final vote between Amsterdam and Milan was a tie, so the result was determined with a coin toss. This has left the Italian side furious — Milan's mayor has even threatened legal action.

The rent issue is adding fuel to this controversy, say agency observers. In February, according to the board-meeting notes, representatives of the Netherlands demanded a 34% increase in the rent for 2019, which had been pegged at €320 (US$390) per square metre of office floor space in the country's initial bid. The increase, they said, resulted from extra costs of fitting out the building and from rising property values in Amsterdam's fast-growing Zuidas neighbourhood, where the EMA's new headquarters are under construction.

The EMA's board rejected the proposal at a 28 February meeting, arguing that the original bid, on the basis of which Amsterdam was chosen, mentioned no such increases and included the costs of fitting out and furnishing the building.

"We have always said that the figures in the initial bid had to be adhered to," says Wathion. "We had a very open and frank discussion with the Dutch colleagues and we reached an agreement."

This final deal includes a clause under which the rent will go up by a fixed rate of 2% every year for the next 20 years, starting at the original price tag of €320 per square metre in 2019 and ending up at €466 per square metre in 2039, according to Wathion. As a result, the EMA's annual rent will increase from an estimated €10 million a year when it moves in, to around €15 million 20 years later.

Pre-agreed rent increases are common in commercial properties, says Wathion, but the fixed increase raised eyebrows at the agency. The meeting minutes called the arrangement "unusual". Rents are typically tied to inflation or the consumer price index. The EMA's London rent, for example, fell by about 4.4% between 2016 and 2018.

But Dutch representatives argued that fixing the rent increase is fair, considering that the price tag, over 20 years, will be smaller than the rent on the EMA's current premises in London's expensive Docklands area. That rent stood at €14.5 million for 2018. The agency's presence will boost property prices in Zuidas owing to an expected influx of professionals, said the Dutch representatives. "The relocation of this agency will work like a magnet for all kinds of companies and professionals," says Anton van Tuyl, a spokesperson for the Dutch Association of Innovative Medicines in The Hague.

The Dutch government had not responded to queries from *Nature* as this story went to press. ∎

The Trace Gas Orbiter in a clean room before its launch in 2016.

PLANETARY SCIENCE

Probe tackles Mars methane mystery

European–Russian Trace Gas Orbiter aims to find source of the gas, a signature of life, on the red planet.

BY NISHA GAIND

For more than a year, a 3.5-tonne spacecraft has been circling Mars in a series of erratic loops. Now, after 1,000 circuits, the Trace Gas Orbiter (TGO) has reached the ideal position for studying the planet's atmosphere, and has made its first scientific observations. It is poised to solve one of the most controversial mysteries in Martian science: why methane, a possible signature of life, is being released on the red planet.

The TGO is part of the ExoMars mission, a joint venture between the European Space Agency (ESA) and the Russian space agency, Roscosmos. The orbiter, which launched in March 2016 and reached the planet that October, is the first craft designed to study gases, including methane, water vapour and ozone, that make up less than 1% of the planet's cold, arid atmosphere. The TGO reached its planned circular orbit earlier this month, and

underwent a series of engineering tests before starting to take scientific data on 21 April.

The Martian atmosphere is composed almost entirely of carbon dioxide, but researchers are interested in its other components — especially methane. These trace gases could be signals of biological or geological activity, and they provide clues to the planet's climatic history.

Scientists have been catching hints of methane in the Martian atmosphere for 15 years using Earth-based telescopes, Mars orbiters and NASA's Curiosity rover. As evidence of the gas has accumulated, the debate over its origin has intensified. "Nearly 95% of all the methane in the Earth's atmosphere originated from current and past biology," says Sushil Atreya, a planetary scientist at the University of Michigan in Ann Arbor. "So, it is natural to ask whether methane on Mars is also of biologic origin."

The gas's presence on Mars has surprised researchers because chemical reactions in the atmosphere should destroy any methane

molecules there within a few centuries. A measurable methane level suggests that an active source must be replenishing the gas.

Curiosity now routinely detects a background level averaging 0.5 parts per billion (p.p.b.) of methane in the atmosphere (by contrast, Earth's level is almost 1,900 p.p.b.). But the Martian concentration changes unexpectedly over time, says Atreya. Observations have also suggested the presence of large plumes with concentrations of 45 p.p.b., and Curiosity has detected burps of around 7 p.p.b. that dissipate quickly. "We didn't expect methane to be on Mars, and it shouldn't be variable unless there is an active source and a vigorous sink," says Bethany Ehlmann, a planetary scientist at the California Institute of Technology in Pasadena.

Researchers have suggested several possible sources. The peaks might come from subsurface chemical reactions between rocks and water, from carbon-rich meteorites that enter the atmosphere or from sudden releases from reservoirs beneath Mars' surface. Most thrilling of all, the peaks could have a biological origin. "It's really a mystery," says Ehlmann.

SCIENTIFIC PAYLOAD

The TGO carries four major scientific instruments that will collect data to create global maps of the trace gases, showing how their presence varies around the planet and over time.

Its main spectrometer, NOMAD, will be able to look down from the craft's orbital altitude through the atmosphere to the surface, a distance of about 400 kilometres. And twice per orbit, when the TGO is on the boundary between Mars's light and dark sides, NOMAD will be able to peer through the planet's atmosphere towards the Sun. This position increases sensitivity by allowing the probe to look through a much bigger segment of atmosphere. Researchers should be able to use the data it collects to identify the spectral signatures of various gases, which absorb light at different wavelengths.

The craft's colour camera, CaSSIS, will build detailed 3D maps of the planet's terrain. And if other instruments pinpoint a spike in methane concentration, researchers will be able to look for images taken at roughly the same time, to try to identify possible sources on the Martian surface. "We're looking for things like an impact crater or a rock slide or landfall, or a big crack in the ground — that would be really nice," says Manish Patel, a planetary scientist at the Open University in Milton Keynes, UK, who works on NOMAD and CaSSIS.

The next major step, says Patel, is to find out whether the instruments are as sensitive as hoped; this will depend on factors such as temperature and dust levels in the atmosphere.

In the first few months, Patel expects to be able to map out levels of methane and start providing the clues that researchers have craved. "The debate over all of this that's been going on for over a decade — it should answer that question," he says. ■

GENETICS

Flu virus finally sequenced in its native form

Technique could help to unpick role of enigmatic chemical modifications in genetic material.

BY EWEN CALLAWAY

The genome of the flu virus has been fully sequenced in its native RNA form for the first time. Previously, all influenza genomes — as well as those of other viruses that store their genetic material as RNA — had been determined by copying the molecule into DNA. The native flu genome was generated using 'nanopore' sequencing technology, which reads RNA strands as they stream through a tiny molecular channel.

"For the first time, we can really start to look at the nature of the genome in its original state," says John Barnes, a microbiologist at the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, who led the effort, described in a preprint posted[1] to the bioRxiv server on 12 April. "It really does start to open up a lot of possibilities."

Barnes and other scientists are most interested in probing viral genomes, as well as other forms of RNA in many types of organism, including humans. Researchers want to investigate mysterious chemical ornaments on RNA molecules that might affect their function in cells but have been hard to study. "The real excitement here is about RNA modifications," says Ewan Birney, co-director of the European Bioinformatics Institute in Hinxton, UK. The approach is "transformative", says Birney.

RNA is chemically similar to its better-known cousin, DNA. In cellular organisms, it serves as an intermediary between DNA-encoded genes and proteins, and performs other tasks. But many viruses — including those behind polio, Ebola and the common cold — store their genetic information as RNA. Barnes, head of the CDC's influenza-genomics team, says that no one had sequenced the virus's RNA genome before because it seemed nearly impossible. Previous methods for sequencing native RNA strands involved degrading one chemical base, or letter, at a time, and these techniques have changed little since their invention in the late 1970s[2]. To compensate, nearly all 'RNA sequencing' instead uses a viral enzyme called reverse transcriptase, which copies RNA into sequencer-friendly DNA strands.

Nanopores offer a simpler way of sequencing actual RNA molecules, such as viral genomes. The technology is based on applying electrical current across a nanoscale molecular pore, and then measuring telltale current fluctuations as genetic material snakes through.
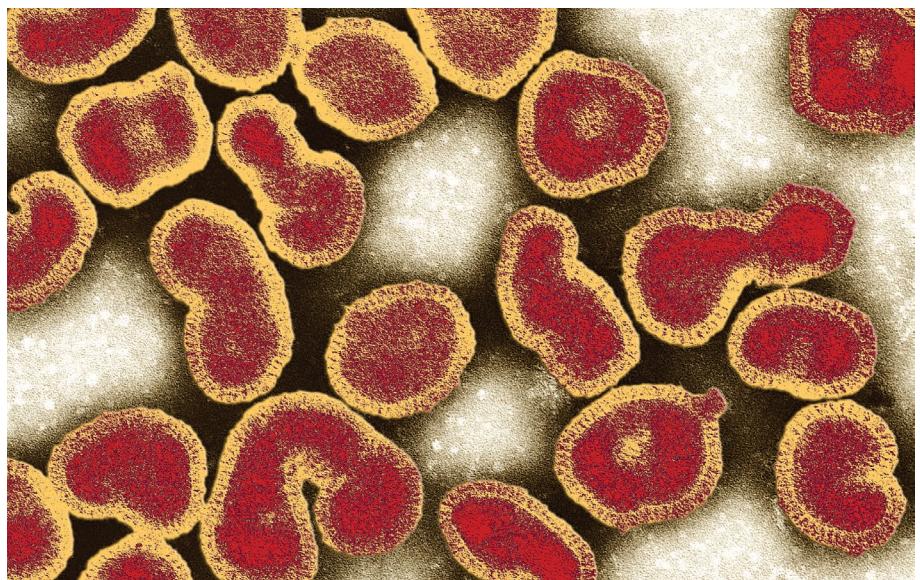
TINY IS MIGHTY

In January, researchers at a leading developer of the technology, UK-based Oxford Nanopore Technologies, directly sequenced RNA using a small device called the MinION[3]. That effort looked at transcripts of messenger RNA, the family of RNA molecules that conveys information from DNA to build proteins.

"It really does start to open up a lot of possibilities."

Barnes's team applied this method to the genome of influenza A, which is roughly 13,500 RNA letters long and composed of 8 segments. Barnes says his team's approach isn't ready for public use. It required a lot of flu virus and, to iron out sequencing errors, the raw data had to be processed many times. But nanopore technology is advancing quickly, and Barnes hopes that with improvements, direct sequencing of RNA viruses will become routine.

At the top of his and other scientists' wish lists are methods for identifying chemical modifications to RNA. More than 100 have been identified so far, but researchers have little idea what most of them do, in large part because it has been impossible to study them systematically. The Oxford Nanopore team was able to detect two common RNA modifications, or tags. Birney, who is a paid consultant to the company, expects that the technology will be able to find many more once machine-learning algorithms are used to unpick the tags' signatures.

Sequencing modified bases of RNA would be "a big deal" for the field, says Bryan Cullen, a virologist at Duke University in Durham, North Carolina. His team found last year[4] that during infection in mice, a tag called m[6]A seems to alter the expression of influenza genes to promote viral replication. Current methods for detecting such modifications are time-consuming and expensive, he adds.

Although the methods aren't yet perfect, Birney says, biologists are still excited about the possibility of soon being able to sequence entire viral genomes and other RNA molecules in their natural forms. "Suddenly, we've got the technology to do this. It's kind of amazing." ∎

1. Keller, M. W. *et al.* Preprint at bioRxiv https://doi.org/10.1101/300384 (2018).
2. Peattie, D. A. *Proc. Natl Acad. Sci. USA* **76,** 1760–1764 (1979).
3. Garalde, D. R. *et al. Nature Meth.* **15**, 201–206 (2018).
4. Courtney, D. G. *et al. Cell Host Microbe* **22**, 377–386.e5 (2017).

The influenza virus is the first RNA virus to have been sequenced in its original state.

DR. GOPAL MURTI/GETTY

CARBON'S FUTURE IN BLACK AND WHITE

THERE ARE REASONS TO BE OPTIMISTIC THAT THE WORLD WILL BREAK ITS ADDICTION TO FOSSIL FUELS. BUT TIME IS RUNNING OUT.

BY JEFF TOLLEFSON

Making sense of recent energy trends can seem like a high-stakes Rorschach test. Some experts see the boom in renewable energy and the shift away from coal in many countries as evidence that the world is beginning to turn a corner on global warming. Others see simply a continuing reliance on low-cost fossil fuels, slow governmental action and a rising risk of planetary meltdown.

The fact is that both sides are right. Renewable energy is indeed undergoing a revolution, as prices for things such as solar panels, wind turbines and lithium-ion batteries continue to plummet. And yet it is also true that the world remains dependent on fossil fuels — so much so that even small economic shifts can quickly overwhelm the gains made with clean energy.

So it was in 2017, when, after staying relatively flat from 2014 to 2016, carbon emissions grew by about 1.5% (see 'A brief lull'). All it took to create that spike was a small rise in economic growth across the developing world, according to a final estimate released in March by the Global Carbon Project, an international research consortium that monitors carbon emissions and climate trends.

The setback is likely to loom large in December, when countries will meet in Katowice, Poland, to complete the

A worker walks through a sorting area at a coal mine in Shanxi, China.

first assessment of their progress in implementing the 2015 Paris climate agreement — an ambitious pact that aims to limit global warming to 1.5–2 °C above pre-industrial levels. The bottom line is hardly encouraging: by and large, governments are falling well short of their commitments, both collectively and individually. Many countries are likely to miss the emissions targets that they made in 2015, and the world is on track for more than 3 °C of warming by the end of the century.

The spike in 2017 threw the situation into sharp relief. Governments are going to have to face the fact that they need to do more if they are serious about meeting the Paris climate agreement's goals, says Glen Peters, a climate-policy researcher at the Center for International Climate Research in Oslo and co-author of the Global Carbon Project's March report. "A lot of hard truths will have to come out in 2018," he says.

Here, *Nature* examines the forces behind the recent emissions trends and what they signal for the future. The good news is that clean-energy technology is at last making substantial strides. The bad news is that the pace isn't nearly quick enough. Big economic and political hurdles stand in the way of shutting off the fossil-fuel spigot and the cheap energy it provides.

THE PLATEAU AND THE SPIKE

To determine where carbon emissions are heading, researchers must first understand why they flattened out for three years. The most optimistic answer is that the seeds of a clean-energy revolution have been planted and are now growing like weeds (see 'The road ahead').

More than a decade of government mandates and economic incentives have helped the renewable-energy industry to take root. Thanks to a combination of technological advances and economies of scale, prices have fallen dramatically for wind and solar. Meanwhile, improvements in lithium-ion batteries have made electric vehicles the clean technology to beat in the transport sector.

All of this has created a virtuous cycle that is driving prices down and sales up, says Jules Kortenhorst, chief executive of the Rocky Mountain Institute, an environmental think tank based in Basalt, Colorado. "President Trump can easily imagine a world where we all turned back to coal-fired electricity and horses and buggies and kerosene lamps, but the reality is that the world is shifting at an accelerating pace to a completely different reality," Kortenhorst says.

The impact of the renewables boom can be readily seen in the United States and China, the world's two largest greenhouse-gas emitters. In the United States, where annual carbon emissions have decreased more than 13% since 2005, renewable sources have become an increasingly important part of the story, contributing more than half of the energy-generating capacity added in 2017 — the equivalent of about 46 average-sized coal plants. In China, the development of renewable energy sources has helped to scale back coal consumption and rein in the country's skyrocketing emissions. In late 2017, Climate Action Tracker, a research consortium that monitors international climate policies, reduced its projection for China's annual emissions in 2030 by 700 million tonnes of CO_2. That figure, which is more than twice the current annual carbon emissions from France, could double if China's efforts to curb coal use continue apace.

But the 2014–16 emissions plateau was shaped by more than just a clean-energy push. One of the biggest factors in keeping levels in check was an economic slowdown in China, which lowered demand for everything from energy to concrete and steel. In the power sector, the country's aggressive pursuit of renewables and nuclear energy has been accompanied by efforts to boost the efficiency of modern coal plants and retire old ones. Similarly, much of the decline in US emissions comes from a shift from coal to natural gas, which emits less carbon.

The upshot, says David Victor, a climate-policy specialist at the University of California, San Diego, is that two of the biggest factors in reducing emissions from electricity come from the fossil-fuel sector itself: increasing coal-plant efficiency in China and the expansion of shale gas in the United States. Because so much energy comes from coal, slight fluctuations from year to year can wipe out massive gains in renewables. "When coal sneezes, the whole world knows it," Victor says. "When solar has a revolution, almost nobody notices."

And, to a large extent, that is precisely what happened in 2017. Solar continued to grow at a breakneck pace, but coal consumption in China also rose. A lack of rainfall in parts of China reduced hydropower output, and coal made up the difference. The government had also initiated a stimulus programme towards the end of 2016, aimed at boosting the economy ahead of the Communist party congress in October 2017. All told, China's carbon dioxide emissions, driven by coal consumption, increased by 3.5% in the first half of 2017, according to the Global Carbon Project.

That small blip in China's coal emissions might have been a major contributor to the spike, but developments in other countries also played a part. India's emissions rose faster than expected, owing to stronger economic growth. Thanks to changes in fossil-fuel consumption, emissions in the United States and European Union dropped more slowly in 2017 than in years past. Then there is the rest of the world, whose emissions rose by 2% in 2017, according to the Global Carbon Project's analysis. That includes developing countries, where tapping fossil fuels remains a relatively cheap and easy way of making economic progress.

BENDING THE CURVE

Little time remains for the world to get its emissions under control. The Paris agreement is predicated on a single global carbon budget that countries are collectively using up each year. The longer humanity waits to reduce emissions, the more aggressive future measures will need to be to keep the total under budget.

It is difficult to say exactly how much time is left. Estimates for the maximum amount of carbon that can be emitted if warming is to remain below 1.5 °C, for example, vary widely. There could be 10 or even 15 years of leeway remaining. Or, humanity might have already burned through the total allotment six years ago. Either way, the tight margins have led many researchers to suspect that even the 2 °C Paris target could be out of reach — at least without developing technologies to pull CO_2 out of the atmosphere or artificially cooling Earth by blocking incoming solar radiation.

The amount by which the world will ultimately warm hinges on a key question: how quickly will the emissions curve bend? An optimist might point to the fact that almost all projections for clean energy have proved to be overly conservative. In 2008, for instance, China set a goal of installing 2 gigawatts of solar photovoltaics by 2020. But it is now likely to achieve more than 200 gigawatts, says Jiang Kejun, a senior researcher at China's Energy Research Institute in Beijing. Kejun says that the pattern is likely to

A BRIEF LULL

Global carbon dioxide emissions plateaued from 2014 to 2016, but rose by 1.5% in 2017.

38
GtCO₂*
2017

34

2008

30
* Gigatonnes of carbon dioxide

PLOTTING THE FUTURE
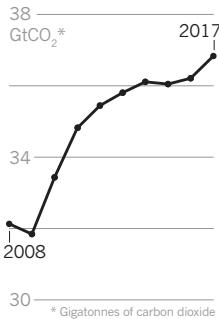
150 GtCO₂* equivalent

Greenhouse-gas emissions could take many paths in the coming years, resulting in differing levels of warming relative to pre-industrial levels. Thanks to policies that have already been implemented by governments around the world, temperatures are not expected to rise as high by 2100 as they otherwise would. But to achieve the 1.5 °C and 2 °C targets set by the 2015 Paris climate accord, more-aggressive emissions reductions will be needed.

100

NO CLIMATE POLICIES
4.1 – 4.8 °C

CURRENT CLIMATE POLICIES
3.1 – 3.7 °C

PLEDGES
2.6 – 3.2 °C

50

▶ Projections

2 °C PATHWAYS

1.5 °C PATHWAYS

0

1990   2000   2020   2040   2060   2080
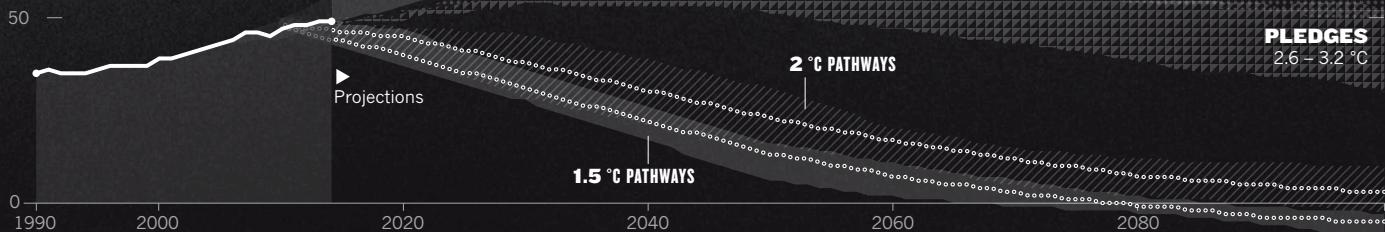
THE ROAD AHEAD

Massive reductions in global greenhouse-gas emissions — of which carbon dioxide is the largest component — will be needed to limit warming to 1.5 or 2 °C. With prices plummeting, clean-energy technologies such as solar, wind and electric vehicles could make a significant dent. But the sheer scale of global fossil-fuel consumption raises questions about the feasibility of that goal.

THE BIG CONTRIBUTORS

A fairly small number of countries are responsible for the bulk of CO_2 released annually. But emissions from the rest of the world are on the rise.

⋯⋯ India   – – European Union
–·– United States   ⋯⋯ China   —— All others

15 GtCO₂*

0

1959   2017

SOLAR TIPPING POINT

In the coming decade or so, the cost of building a solar plant is expected to fall below the cost of operating an existing coal plant in Germany, China and other countries.

—— Coal   ⋯⋯ Gas
⋯⋯ Wind   —— Solar

80 US$ per megawatt hour

0

2017   **Germany**   2040   2017   **China**   2040
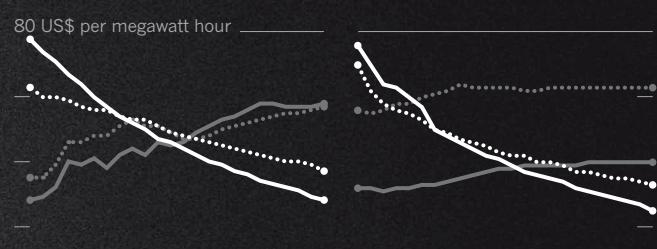
THE SCALE OF THINGS

Global energy consumption is still dominated by fossil fuels. A small fluctuation in coal use from one year to the next can wipe out a seemingly dramatic expansion in renewable energy.

* Gigatonnes of carbon dioxide
** Gigatonnes oil equivalent

5 Gtoe**

2016

2000

0

Coal   Oil   Gas   Nuclear   Hydro   Other renewables

SEEDS OF A REVOLUTION

The cost of electricity generated by solar modules has declined significantly in the past 20 years.

7 US$ per watt

0

1998   2017
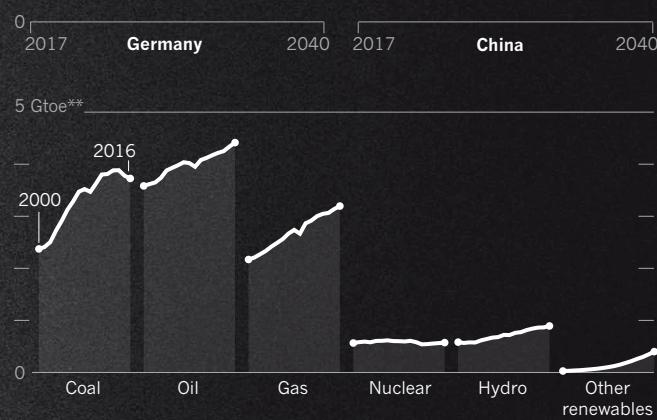
be repeated in the future. "Modellers are underestimating the potential of renewable energy," he says.

Some analysts think that solar energy, in particular, is poised to hit a tipping point that could change the face of the energy market. Watt for watt, solar energy already costs as little as coal in some places. And intriguingly, the London-based energy consultancy Bloomberg New Energy Finance (BNEF) has calculated that solar could become so cheap that, by 2030, it would be more cost-effective in many regions to build a solar plant than to continue supplying fuel to an existing coal plant. Similarly, beginning in the mid-2020s, the consultancy projects that falling battery prices will make electric cars cheaper to buy and run than their conventional counterparts — without the government subsidies that have fuelled the market so far.

"These are very important tipping points, but nobody quite knows what will happen and how policymakers will respond," says Angus McCrone, chief editor at BNEF. "Politics is a bottleneck, particularly when there's a powerful incumbency in an industry that is going to be affected by these new technologies."

But politics can also help to bring about rapid change. While Trump is fighting on behalf of the fossil-fuel industry, leaders of other countries are moving in the opposite direction. The United Kingdom and France have both announced plans to ban the sale of petrol- and diesel-powered vehicles by 2040. And more than two dozen countries have committed to phasing out coal by as early as 2030.

These types of mandate are a sign that energy politics might be shifting towards more brute-force methods, says Michael Mehling, an energy and environmental-policy researcher at the Massachusetts Institute of Technology in Cambridge. Economists tend to favour market-based programmes, such as the EU's Emissions Trading System, but Mehling says there is little evidence that such arrangements will drive the kind of rapid transformational change needed to meet global climate goals. Old-school government mandates might be the last resort, Mehling says.

"If the decisions are made at a sufficiently high level," he says, "they can change the landscape pretty much overnight".

Kejun's calculations suggest that, driven both by policy and economics, China's carbon emissions are still on track to peak as early as 2020, and its coal consumption could drop by as much as 40–50% by 2030. "The transition has already started," says Kejun.

A similar movement seems to be under way in India, which is racing to provide reliable power — and cleaner air — to more than 1.3 billion people. If India can chart a path to sustainable development, it would set an example for other developing countries and avoid a repeat of China's coal-fuelled ascension.

Today, the solar-power industry is booming in India, thanks to government incentives and falling prices, and the Indian government aims to install 100 gigawatts of solar capacity by 2022 — nearly double the current solar-generation capacity in the United States. Meeting that goal could be challenging, because solar power will increasingly need to compete with existing coal-fired power plants for limited space on the electricity grid, says Rahul Tongia, an energy researcher at the non-profit public-policy organization the Brookings Institution in New Delhi. Still, he says, the trends are impressive. "Maybe it takes a bit longer to hit the targets. Who cares?" Tongia says. "The progress is still remarkable, measurable, dramatic and meaningful."

But can such progress realistically rein in warming? For Peters, the boom in renewable energy is necessary and welcome, but still insufficient. Ultimately, the only thing that matters to the climate is the quantity of greenhouse gases emitted — and so the question is when humanity will begin to close the spigot and shut down fossil-fuel infrastructure. When that happens, he says, "you can start to feel a little bit better". ∎

Work is done on a rooftop solar panel installation in Wuhan, China.

Jeff Tollefson writes for *Nature* from New York City.
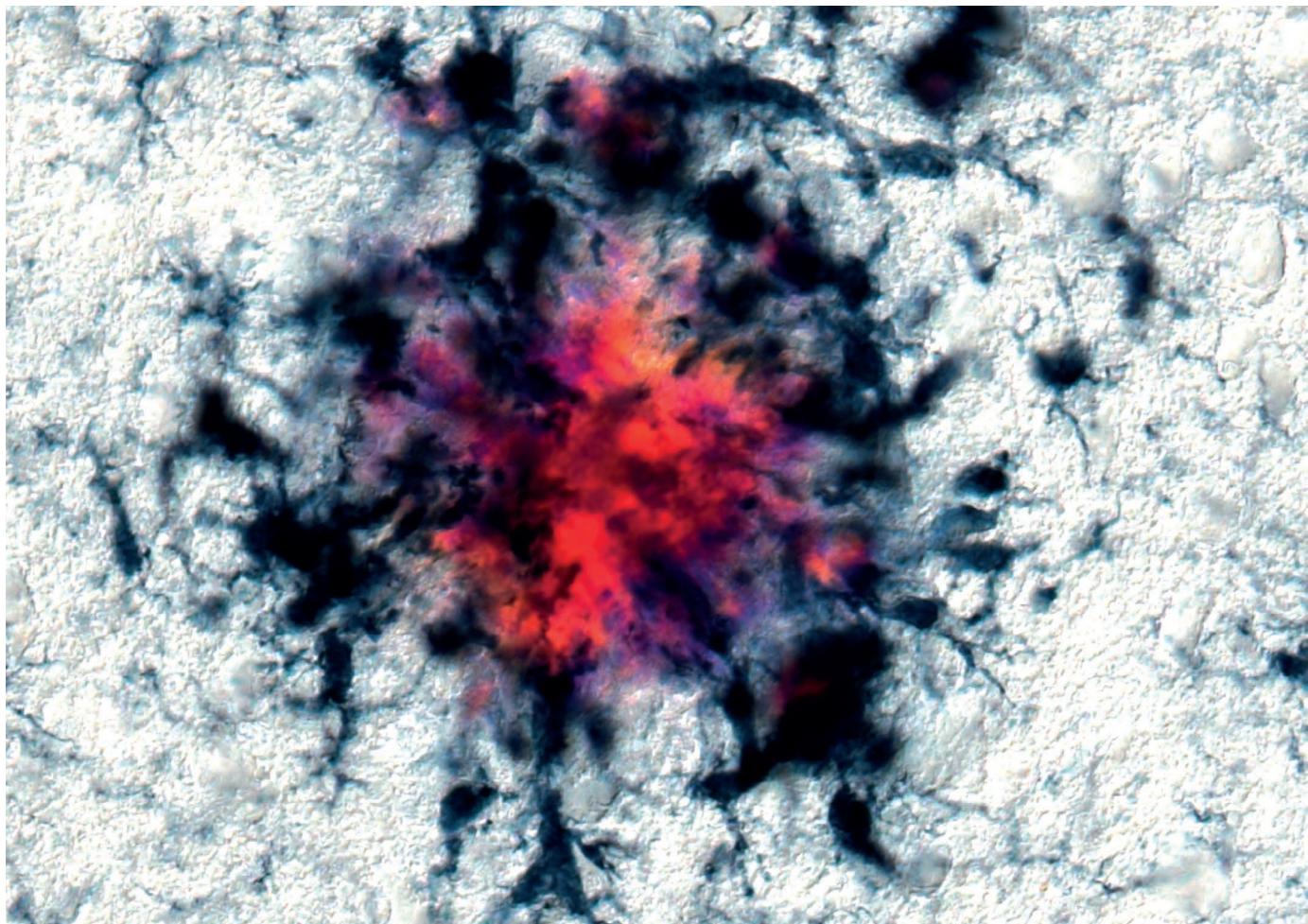
THE BRAIN INFLAMED

The brain's immune system could be provoking Alzheimer's and other neurodegenerative diseases. Can scientists get it back in check?

BY ALISON ABBOTT

Neuroscientist Michael Heneka knows that radical ideas require convincing data. In 2010, very few colleagues shared his belief that the brain's immune system has a crucial role in dementia. So in May of that year, when a batch of new results provided the strongest evidence he had yet seen for his theory, he wanted to be excited, but instead felt nervous.

He and his team had eliminated a key inflammation gene from a strain of mouse that usually develops symptoms of Alzheimer's disease. The modified mice seemed perfectly healthy. They sailed through memory tests and showed barely a sign of the sticky protein plaques that are a hallmark of the disease.

Yet Heneka knew that his colleagues would consider the results too good to be true.

Even he was surprised how well the mice fared; he had expected that removal of the gene, known as *Nlpr3*, would protect their brains a little, but not that it would come close to preventing dementia symptoms. "I thought something must have gone wrong with the experiments," says Heneka, from the German Center for Neurodegenerative Diseases in Bonn.

He reanalysed the results again and again. It was past midnight when he finally conceded that they might actually be true.

Over the next couple of years, he confirmed that nothing had gone wrong with the experiments. Together with his colleagues, he replicated and elaborated on the results[1]. Since then, numerous studies have bolstered the link between dementia and the brain's immune system, highlighting the cells and signals involved[2]. But none has managed to fully pin it down — the link seems to be slippery and dynamic, changing as the disease progresses.

Even so, the idea has sparked the interest of pharmaceutical investors, who see a large, and entirely unserved, market: an estimated 50 million people worldwide have dementia — a number the World Health Organization projects will rise to 82 million by 2030. Of the eight drug-discovery projects backed by Dementia Consortium — a UK-based group of charities and pharmaceutical companies that has

Microglia cluster around plaques in a mouse model of Alzheimer's disease.

poured £4.5 million (US$5.7 million) into the projects — four are aimed at inflammation.

But there are roadblocks ahead. Scientists don't yet agree on whether the immune system will need to be ramped up or tamped down at different stages of disease. And some of the practical problems that have dogged clinical trials in Alzheimer's disease — imperfect mouse models and difficulties in recruiting patients early enough — may plague this new approach, too. Hanging over the field like a black cloud is the fact that all clinical trials in Alzheimer's disease have so far failed.

Still, bioinformatician Martin Hofmann-Apitius at the Fraunhofer Institute for Algorithms and Scientific Computing in Sankt Augustin, Germany, who specializes in pharmaceutical research, notes that researchers have filed several patents relating to inflammation-related targets. "Soon we will see a wave of clinical trials," he predicts.

CLOGGED AND SWOLLEN

The German psychiatrist Alois Alzheimer was the first to described the symptoms and pathology of dementia, in the early twentieth century. Looking under the microscope at the brain of a woman whose cognitive decline he had witnessed, he saw — and neatly drew — the plaques, now known to contain amyloid-β, and tangles of a protein called tau that together are the signature of the disease. In those earliest depictions of the affected brain tissue, Alzheimer also sketched microglia, a type of immune cell in the brain, nestling next to neurons. "Alzheimer himself noticed the cells and drew them in abundant number alongside neurons," says Heneka.

Although the sketches made no deeper link between microglia and disease, Heneka remembered them as links between inflammation and Alzheimer's began to emerge in the mid-1990s. He had been intrigued by some epidemiological observations showing that people given some anti-inflammatory drugs (to treat rheumatoid arthritis, for instance) seemed to be at a lower risk of developing Alzheimer's disease than the general population. He became encouraged by reports that microglia gather around plaques and areas of brain degeneration, and that inflammatory molecules such as cytokines collect in the cerebrospinal fluid of patients. Most scientists assumed that these observations reflected a passive response to tissue damage. But Heneka always suspected that inflammation could be actively provoking disease.

Microglia have turned out to be central to the link between inflammation and neurodegeneration (see 'Help or hinder'). The cells have two major functions. They take care of the general health of neurons and their synapses — the junctions between neurons where they communicate with one another. And they patrol the brain, searching for threats and problems. When

they detect an infectious or otherwise-aberrant molecule such as amyloid-β — or debris from damaged cells — they become activated and signal to other microglia to join them in a clean-up effort. Certain microglial proteins gather into large complexes called inflammasomes (a key component of the inflammasome is Heneka's NLRP3 protein), which churn out clean-up signals in the form of activated immune molecules. Inflammasomes usually ebb away once the job is done, but in Alzheimer's they seem to remain activated, continuing to pump out inflammatory molecules yet failing to clean up properly.

In 2013, microglia began to loom large in Alzheimer's disease research. Around the same time that Heneka's paper showed that preventing inflammation staved off Alzheimer's pathology in mice, the *New England Journal of Medicine* published two large studies of gene variants associated with the disease[3,4]. Both studies linked the risk of developing

"WE JUST DON'T KNOW ENOUGH ABOUT THE BIOLOGY YET."

late-onset Alzheimer's to a gene called *TREM2*, which makes a protein that sits in the membrane of microglial cells.

Neuroscientists started to pay attention. So did immunologists. An interdisciplinary community of neuroimmunologists burgeoned. "Suddenly, huge opportunities opened up," says neuroscientist Michela Matteoli at the University of Milan, Italy, who now runs a neuroscience programme in the immunology department at the neighbouring Humanitas Institute. At Humanitas, she found a treasure trove of mouse models lacking specific elements of the immune system, which immunologists had never had reason to use for studying brain function. "Many of the tools we need are available," she says.

HEROES AND VILLAINS

How might microglia, which evolved to keep the brain in good order, become a force for the bad in Alzheimer's? Last year, Heneka and his colleagues published evidence suggesting a plausible mechanism for the switch, at least in their mice. They found that activated microglia discard the remnants of inflammasomes in tiny clumps called specks, and that these specks go on to seed new amyloid-β clusters, spreading the disease across the brain[5]. "A perfect storm," says Heneka. "Toxic amyloid-β promotes inflammation, which promotes more toxic amyloid-β."

He is working together with immunologist Eicke Latz, at the University of Bonn, to develop a drug that can stop the inflammasome from forming. That would allow the microglia to continue their other important roles in the brain's housekeeping without conscripting other microglia to help clean up. The storm would be kept at bay.

Latz co-founded the start-up IFM Therapeutics in Boston, Massachusetts, in 2016. The company, which was acquired by the pharmaceutical firm Bristol Myers Squibb last year, already has some candidate drugs that stop inflammasomes from forming, and Latz and Heneka hope to start clinical trials in the next couple of years.

Meanwhile, neuroimmunologists around the world are trying gain a deeper understanding of the biology of microglia, to work out whether there could be other ways to design immune-based therapies for Alzheimer's and other neurodegenerative diseases. Some scientists think that the healthy activities of microglia could be bolstered to clear toxic amyloid-β more efficiently and avoid the storm altogether.

Two studies in mice and post-mortem human brains have shown that the microglia that huddle around plaques in the brain are a very specific subset[6,7]. They express some genes at higher or lower levels than regular microglia, and those patterns tell an interesting story: the cells seem to be trying to tune up their normal housekeeping duties to combat the plaques. Some of those genes remove safeguards, or 'check-points', from the pathways that lead to the cells' activation. Others are in pathways that sense damage or encourage microglia to engulf defective molecules. In each case, the gene-expression patterns indicate that the microglia are ramping up their house-keeping duties to try to protect the brain.

Mutations in about a dozen of these genes had already been identified as risk factors for Alzheimer's in humans, says Ido Amit, an immunogeneticist at the Weizmann Institute of Science in Rehovot, Israel, who conducted one of the studies looking at the gene-expression patterns[6].

Amit says that the cells are clearly there for a reason and might therefore be harnessed to help. "The results seemed to be telling us a strong message about the biology of the system," he says. If microglia could be helped to perform their regular functions more efficiently, and kept from any overzealous cleaning efforts, they might help stave off symptoms of the disease rather than worsening its course.

If there were any doubts still lingering about the importance of microglia in mechanisms of dementia — whether they serve as heroes or villains — these papers eliminated them. What's more, microglia could even be primed for activation by inflammation elsewhere in the body. Epidemiological studies have shown that the burden of infection

during life increases the risk of cognitive impairment or dementia in later life[8,9]. And earlier this month, Jonas Neher from the German Center for Neurodegenerative Diseases in Tübingen and his colleagues showed that provoking inflammation in mice by injecting molecules called lipopolysaccharides (LPS) into their bellies led to persistent changes in gene expression in brain microglia — even though the molecules themselves didn't enter their brains. Low doses of LPS led to increased levels of amyloid-β and plaques; high doses reduced the burden[10].

Microglia could even be involved in other neurodegenerative diseases, because similar findings have been observed in models of amyotrophic lateral sclerosis (ALS) and Parkinson's disease[11]. And research from Matteoli and others suggests they could be implicated even more widely in brain disorders, such as the rare neurodevelopmental disorder known as Rett syndrome[12,13].

SHELTER FROM THE STORM
Amit is now discussing with industrial partners how the housekeeping activities of microglia might be boosted. "This would allow us to reactivate our natural defences when damage is out of control," he says.

Others worry that activating more microglia in late stages of the disease might make things worse. "We just don't know enough about the biology yet," says Oleg Butovsky, a neuroimmunologist at Harvard Medical School in Boston, who led the other study on gene expression in microglia[7] and is developing biomarkers to identify them in the brain at different stages of the disorder. He says it isn't clear whether microglia should be boosted or suppressed, or even whether different tactics could be used at different times during the progression of the disease.

And not all scientists assume that the role of the immune system in neurodegeneration stops with microglia. Neurologist Philip De Jager at Columbia University in New York is developing an Alzheimer's therapy that is based on a microglial target, but says that cells from the rest of the body's immune system, such as T cells, which are present in very low numbers in the brain, might also turn out to be relevant.

Although clinical interest is taking off, there

HEALTHY NEURON

HELP OR HINDER
Microglia are the brain's housekeeping squad, clearing away damaged cells and their components to keep neurons healthy. They can overreact to the build-up of protein plaques associated with Alzheimer's disease.

Amyloid plaque

MICROGLIA IN DISEASE

NORMAL MICROGLIA FUNCTION

Try to fight amyloid plaque build-up

Overproduce inflammatory signals

Flag and consume myelin if it becomes damaged

Myelin

Release proteins that seed new plaques

Synapse

Digest infectious or faulty cells or molecules

Monitor messages between neurons

Wrongly mark synapses for destruction, worsening neurodegeneration

NIK SPENCER/*NATURE*

are two stubborn elephants in the room: the mouse models used in Alzheimer's research are a poor proxy for the human condition, and it is difficult to find people who are good candidates for testing new therapies.

Mice with gene mutations that predispose them to Alzheimer's develop some realistic symptoms, but too quickly. That leaves scientists struggling to identify when treatment should be given. "Our models are just too accelerated," says Marco Colonna from Washington University School of Medicine in St. Louis, who has worked extensively on the biology of TREM2. "The field recognizes that the development of a model where amyloid accumulates more naturally is a priority."

It's also a challenge to identify people early enough in the progression of their disease for any experimental drug to have a chance of working. Alzheimer's researchers think that many of the earlier trials failed not because their hypothesis — that amyloid-β and tau

are critically involved in the disease — is incorrect, but because the treatment is given too late. Patients are generally recruited to trials only after their plaque burden and neurodegeneration has advanced and the disease is probably irreversible. This could also be one reason why trials of anti-inflammatory drugs such as naproxen or rofecoxib have gone the same way as other potential treatments and shown no benefit in people with Alzheimer's, says Heneka. Biomarkers to identify people who are in a very early stage of disease are only now becoming available. Even then, the tests are very expensive and cumbersome, involving brain scans and spinal taps. And they still need to be completely validated in practice.

The many uncertainties are not damping enthusiasm. "It's been an exciting few years," says De Jager. Scientists in the field see a parallel with cancer immunotherapy, where the immune system receives a boost to attack tumours. "It seems that diseases not thought classically to be immunological may indeed have an immunological basis."

When Heneka thinks back to his experiments with the unexpectedly smart mice, he is cautiously optimistic that immune-based therapies could work for Alzheimer's disease. But the new trials need to face up to the troubles that plagued previous efforts. No one, he says, wants to see the approach fail for the wrong reasons. Then again, he had never seen a mouse that was supposed to have Alzheimer's pass a memory test with such flying colours. ■

Alison Abbott *is* Nature's *senior European correspondent.*

1. Heneka, M. T. *et al. Nature* **493**, 674–678 (2013).
2. Salter, M. W. & Stevens, B. *Nature Med.* **9**, 1018-1027 (2017).
3. Jonsson, T. *et al. N. Engl. J. Med.* **368**, 107–116 (2013).
4. Guerreiro, R. *et al. N. Engl. J. Med.* **368**, 117–127 (2013).
5. Venegas, C. *et al. Nature* **552**, 355–361 (2017).
6. Keren-Shaul, H. *et al. Cell* **169**, 1276–1290 (2017).
7. Krasemann, S. *et al. Immunity* **47**, 566–581 (2017).
8. Wallin, K. *et al. J. Alzheimer's Res.* **31**, 669-676 (2012).
9. Bu, X.-L. *et al. Eur. J. Neurol.* **22**, 1519–1525 (2015). 10.1111/ene.12477
10. Wendeln, A.-C. *et al. Nature* https://doi.org/10.1038/s41586-018-0023-4 (2018).
11. Yeh, F.L. *et al. Trends Mol. Med.* **23**, 512-533 (2017).
12. Tomasoni, R. *et al. eLife* **6**, e21735 (2017).
13. Derecki, N. C. *et al. Nature* **484**, 105–109 (2012).

COMMENT

Simplified 3D brain organoids can be grown in a dish using human stem cells as the starting material.

The ethics of experimenting with human brain tissue

Difficult questions will be raised as models of the human brain get closer to replicating its functions, explain **Nita A. Farahany**, **Henry T. Greely** and 15 colleagues.

If researchers could create brain tissue in the laboratory that might appear to have conscious experiences or subjective phenomenal states, would that tissue deserve any of the protections routinely given to human or animal research subjects?

This question might seem outlandish. Certainly, today's experimental models are far from having such capabilities. But various models are now being developed to better understand the human brain, including miniaturized, simplified versions of brain tissue grown in a dish from stem cells — brain organoids[1,2]. And advances keep being made.

These models could provide a much more accurate representation of normal and abnormal human brain function and development than animal models can (although animal models will remain useful for many goals). In fact, the promise of

brain surrogates is such that abandoning them seems itself unethical, given the vast amount of human suffering caused by neurological and psychiatric disorders, and given that most therapies for these diseases developed in animal models fail to work in people. Yet the closer the proxy gets to a functioning human brain, the more ethically problematic it becomes.

There is now a need for clear guidelines for research, albeit ones that can be ▶

GENOME INSTITUTE OF SINGAPORE, A*STAR

▶ adapted to new discoveries. This is the conclusion of many neuroscientists, stem-cell biologists, ethicists and philosophers — ourselves included — who gathered in the past year to explore the ethical dilemmas raised by brain organoids and related neuroscience tools. A workshop was held in May 2017 at the Duke Initiative for Science & Society at Duke University in Durham, North Carolina, with limited support from the US National Institutes of Health (NIH) BRAIN Initiative. A similar US meeting was held last month on related topics.

Here we lay out some of the issues that we think researchers, funders, review boards and the public should discuss as a first step to guiding research on brain surrogates.

SAFE SURROGATES

Three classes of brain surrogate offer researchers a way to investigate how the living human brain works, without the need for potentially risky — if not ethically impossible — procedures in people.

Organoids. Brain organoids can be produced much as other 3D multicellular structures resembling eye, gut, liver, kidney and other human tissues have been built[2–4]. By adding appropriate signalling factors, aggregates of pluripotent stem cells (which have the ability to develop into any cell type) can differentiate and self-organize into structures that resemble certain regions of the human brain[5–7].

Investigators use different approaches. They might coax pluripotent stem cells to turn into specific populations of neural cells, such as those specific to a particular brain region. Or they can allow the pluripotent cells to differentiate on their own, in which case both neural cells and other cell types might be generated[2]. Brain organoids resembling particular brain regions can even be combined into 'brain assembloids' to enable researchers to study the formation of neural circuits and cellular interactions between different regions[8].

Compared with 2D sheets of neural cells in a dish, the 3D structures last longer (for around two years[9]) and can consist of more types of cell. They also mimic key features of developing brains. For instance, in later stages of fetal development, the cerebral cortex switches from generating neurons to creating glial cells (the various other cell types in the brain that nourish, surround and protect neurons). This process can be captured in brain organoids, allowing investigators to gain insights that would be experimentally and ethically extremely challenging, if not ethically unacceptable, to obtain from developing brains.

Already, researchers have deployed brain organoids to investigate neurodevelopmental alterations in people with autism spectrum disorders[8,10] or schizophrenia[11], and to study the unusually small brain size (microcephaly) seen in some babies infected with the Zika virus before birth[12].

Brain organoids have limitations. They lack certain cell types, such as microglia and cells that form blood vessels. Today, the largest organoids are about 4 millimetres in diameter and contain only about 2 million to 3 million cells. An adult human brain measures roughly 1,350 cubic centimetres, and is made up of 86 billion neurons and a similar number of non-neuronal cells. Moreover, so far, brain organoids have received sensory input only in primitive form, and connections from other brain regions are limited.

Given such constraints, the possibility of organoids becoming conscious to some degree, or of acquiring other higher-order properties, such as the ability to feel distress, seems highly remote. But organoids are becoming increasingly complex. Indeed, one of us (P.A.) recorded neural activity from an organoid after shining light on a region where cells of the retina had formed together with cells of the brain. This illustrated that an external stimulus can result in an organoid response[13].

***Ex vivo* brain tissue.** Another type of model involves slices of brain tissue that have been removed from individuals during some surgical procedure, for example to treat seizures.

For more than a century, researchers have studied brain cells in tissue extracted from patients undergoing surgery, or from people who have died. But technological advances, including in imaging and in the techniques used to preserve the functional properties of brain tissues in the lab (*ex vivo*), could make this approach considerably more powerful.

> "The possibility of organoids becoming conscious to some degree seems highly remote."

When tissue from the neocortex or hippocampus regions is removed to treat a pathology, such as epilepsy or cancer, the piece removed is typically the size of a sugar cube (about 1–4 cubic centimetres), although it can sometimes be much bigger. That piece is then generally cut into slices, the functional properties of which can be preserved for weeks.

Using these slices, researchers can measure the synaptic and other properties of neurons in intact brain circuits; map the 3D morphology of circuits; and extract and analyse cellular RNA to probe gene expression. They can also manipulate the firing of specific neurons using optogenetics, which could enable them to analyse in more detail the functional properties of human brain circuits. (Optogenetics uses light to track or selectively activate neurons that have been genetically modified to express a light-sensitive protein.)

Currently, *ex vivo* brain tissue does not have sensory inputs. And with outbound connections severed, isolated tissues can't communicate with other regions of the brain, or generate motor outputs. Thus, the possibility of consciousness or other higher-order perceptive properties emerging seems extremely remote.

Chimaeras. The third class of experimental brain model involves the transplantation of human cells, derived *in vitro* from pluripotent stem cells, into the brains of animals such as rodents. This can be done while the animal fetus is developing or after the animal is born. Such chimaeras are generated to provide a more physiologically natural environment in which the human cells can mature.

Neuroscientists have transplanted human glial cells into mice, for instance, and found that the animals perform better in certain tasks involving learning. Researchers have also injected human stem cells into early-stage pig embryos, and then transferred the embryos into surrogate sows, where they've been allowed to develop until the first trimester. More than 150 of the embryos developed into chimaeras; in these embryos, about 1 in 10,000 cells in the precursors of hearts and livers were human.

In principle, chimaeras could help researchers to better understand human illnesses and the effects of drug treatments. Labs have developed human–mouse chimaeras to shed light on Parkinson's disease, for example.

Some groups have even successfully transplanted human brain organoids into rodents, where they have become supported by blood vessels (vascularized)[14]. The provision of a blood supply is an essential step in enabling organoids to grow larger than their current achievable size. But the size of rodent models restricts the degree to which human brain organoids can grow within them.

ISSUES TO CONSIDER

Currently, if research on human tissue occurs outside a living person, only the processes of obtaining, storing, sharing and identifying the tissue fall under the regulations and guidelines that limit what interventions can be conducted on people. As brain surrogates become larger and more sophisticated, the possibility of them having capabilities akin to human sentience might become less remote. Such capacities could include being able to feel (to some degree) pleasure, pain or distress; being able to store and retrieve memories; or perhaps even having some perception of agency or awareness of self.

Could studies involving brain tissue that has been removed from a living

A researcher dissects slices of human brain tissue.

person or corpse provide information about the person's memories, say? Could organisms that aren't 'biologically human' ever warrant some degree of quasi-human or human moral status?

In the light of such possibilities, here we lay out some of the issues that we think civil society, researchers, ethicists, funders and reviewers ought now to be considering.

Metrics. Is it even possible to assess the sentient capabilities of a brain surrogate? What should researchers measure? If appropriate metrics can be developed, how do investigators decide which capabilities are morally concerning?

Neuroscientists have made considerable progress when it comes to identifying the neural correlates of consciousness[15]. Yet the signals for consciousness or unconsciousness detected in a living adult — using electroencephalography (EEG) electrodes, for example — don't necessarily translate to infants, animals or experimental brain surrogates. Without knowing more about what consciousness is and what building blocks it requires, it might be hard to know what signals to look for in an experimental brain model[15].

With regard to human–animal chimaeras, researchers are already dealing with beings that have some form of consciousness. Here, the need to establish what measures to base protections on (both for

the animal and the human subject) is more pressing. One possibility is for researchers to use anaesthetics or other methods to maintain comatose-like brain states. Perhaps certain brain functions or a pre-specified level of brain activity, signalling a lack of capacity, could be used to delineate ethically justifiable research.

Human–animal blurring. Researchers have already produced mice with rat pancreases by injecting rat pluripotent stem cells into mouse embryos. The same approach could one day enable the production of human organs in other animals[16].

How do we define the boundaries of this research? What implications might such boundaries have for vascularizing brain organoids, or for growing neural tissue in animals? Is the production of a human heart in a pig's body acceptable, for instance, but not the production of a brain from human cells?

We believe that decisions about which kinds of chimaera are permitted, or about whether certain human organs grown in animals make animals 'too human-like', should ultimately be made on a case-by-case basis — taking into account the risks, benefits and people's diverse sensitivities.

Death. Do *ex vivo* human brain models challenge our understanding of life and death? What implications might such models have for the legal definition of

death, and what are the implications for decisions tied to this definition, such as organ donation?

The advent of tracheal positive-pressure ventilation in the 1950s and cardiopulmonary resuscitation (CPR) in the 1960s led to the concept of brain death. Beginning in the 1960s, a person whose brain had completely and irreversibly ceased to function could be declared dead, even if they still had a heartbeat.

Any emerging technologies that could restore lost functionality to a person's brain could potentially undermine the diagnosis of brain death, because the cessation of brain function might no longer be permanent and irreversible. But a distinction here is important: technologies that would restore a few neurons or certain limited kinds of brain activity would not restore clinical functionality of the brain and so would not raise this concern.

Consent. Is the standard process of obtaining informed consent adequate for research using human brain cells or tissue, or developing brain surrogates from induced pluripotent stem cells?

Currently, researchers using pluripotent stem cells or brain tissues generally disclose their plans to donors in broad terms. Given how much people associate their experiences and sense of self with their brains, more transparency and assurances

could be warranted. Donors might wish to deny the use of their stem cells for the creation of, say, human–animal chimaeras.

This targeted approach is used in other contexts. When people undergoing *in vitro* fertilization procedures choose to donate excess embryos to research, for instance, they are assured that these will not be used to create a baby.

Stewardship. Is there a point at which we should be concerned about the welfare of brain surrogates or chimaeras, such that assigning someone loosely akin to a guardian or decision-maker for the brain surrogate might be warranted, beyond the researchers involved? Such an arrangement would be similar to the appointment of a guardian *ad litem* in custody disputes involving children in the United States (someone besides the parents who can represent the child's interests).

Ownership. Who, if anyone, should 'own' *ex vivo* brain tissue, brain organoids or chimaeras?

At present, brain tissue samples are owned by the researchers or organizations collecting the tissue or doing the science. If significant developments in the field one day lead us to regard any of these brain surrogates as having greater moral status than we would currently give them, might greater privileges and protections be appropriate?

Post-research handling. How should human brain tissue be disposed of, or handled at the end of an experiment?

Today, brain organoids or *ex vivo* brain tissue are destroyed following standard practices for disposing of all tissues. But if researchers develop mice, say, with some advanced cognitive capacities, should those animals be destroyed or given special treatment at the end of a study? Already certain animals, such as chimpanzees, enter sanctuaries to live out the remainder of their lives after researchers have finished working with them in laboratories.

Data. Should there be special requirements for data sharing, collaboration and legacy use of brain tissue?

The unique benefits and risks of sharing data obtained from such tissues will need to be considered. *Ex vivo* human brain tissue could reveal sensitive information — for instance, about a person's memories or disease status. Equally, there could be more value in sharing such information, because of the difficulty of obtaining human brain tissue. In some cases, certain features of the data might need to be stripped out, or the extent of sharing limited.

Geneticists have long grappled with similar issues for people's genomic information; some of their approaches could be applied to brain research.

ETHICS EFFORTS

Various efforts are already tackling the ethics of advances in neuroscience[17]. When the BRAIN Initiative was announced in 2013, the Presidential Commission for the Study of Bioethical Issues was charged with evaluating ethics, and produced a two-volume report in response[18,19]. The European Commission's Human Brain Project has a major ethics component, and the NIH BRAIN Initiative has a neuroethics division.

A 3D human-brain assembloid derived from stem cells.
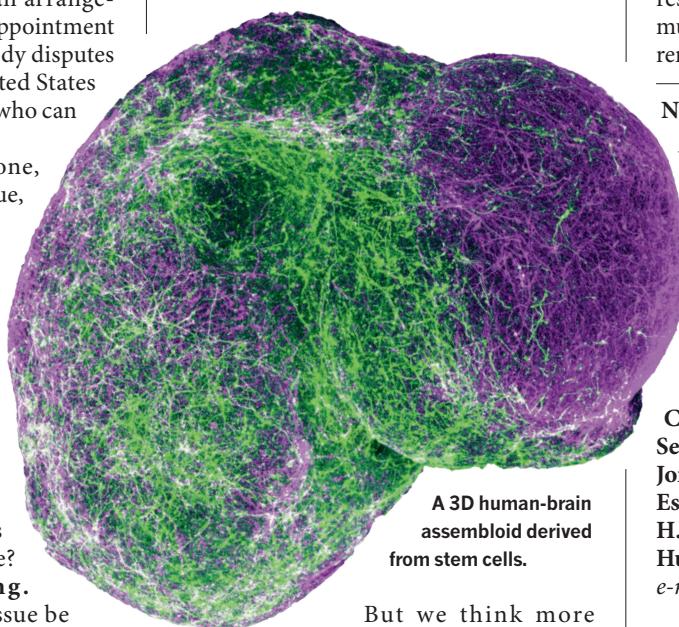
But we think more needs to be done. Existing institutional ethics review boards or those for stem-cell research oversight might not yet be equipped to address issues specific to these experimental brain models because they are so new. We recommend that such organizations ask experts in this area to join their boards or serve as consultants. New committees, dedicated to overseeing the use of human-brain surrogates, could also be assembled.

As for the broader societal conversation, various models exist for democratic deliberation that could be applied. One example is the successful consultations between the public, scientists, regulators and bioethicists that preceded the UK government's decision to permit the clinical use of mitochondrial DNA transfer in 2015.

As these conversations play out, the major funders of biomedical research should strive to provide guidance and,

"Models of the human brain could help us to unlock mysteries about psychiatric illnesses that have long remained elusive."

eventually, guidelines. Also, researchers engaged in the development and use of human-brain surrogates should seek ethical guidance, for instance from their funders, review boards or institutions. They should also share their experiences and concerns, as reviewers, in their own papers or at conferences.

We do not think that these difficult questions should halt this research. Experimental models of the human brain could help us to unlock mysteries about psychiatric and neurological illnesses that have long remained elusive. But to ensure the success and social acceptance of this research long term, an ethical framework must be forged now, while brain surrogates remain in the early stages of development. ∎

Nita A. Farahany *is professor of law and philosophy at Duke University, director of the Duke Initiative for Science & Society, Duke University, Durham, North Carolina, USA.* **Henry T. Greely** *is professor of law, director of the Center for Law and the Biosciences, and director of the Stanford Program in Neuroscience and Society at Stanford University, California, USA.*
Steven Hyman, Christof Koch, Christine Grady, Sergiu P. Paşca, Nenad Sestan, Paola Arlotta, James L. Bernat, Jonathan Ting, Jeantine E. Lunshof, Eswar P. R. Iyer, Insoo Hyun, Beatrice H. Capestany, George M. Church, Hao Huang, Hongjun Song.
e-mail: farahany@duke.edu

1. Quadrato, G., Brown, J. & Arlotta, P. *Nature Med.* **22**, 1220–1228 (2016).
2. Pasca, S. P. *Nature* **553**, 437–445 (2018).
3. Arlotta, P. *Nature Meth.* **15**, 27–29 (2018).
4. Eisenstein, M. *Nature Meth.* **15**, 19–22 (2018).
5. Lancaster, M. A. & Knoblich, J. A. *Nature Protoc.* **9**, 2329–2340 (2014).
6. Qian, X. *et al. Cell* **165**, 1238–1254 (2016).
7. Kadoshima, T. *et al. Proc. Natl Acad. Sci. USA* **110**, 20284–20289 (2013).
8. Birey, F. *et al. Nature* **545**, 54–59 (2017).
9. Sloan, S. A. *et al. Neuron* **95**, 779–790 (2017).
10. Mariani, J. *et al. Cell* **162**, 375–390 (2015).
11. Ye, F. *et al. Neuron* **96**, 1041–1054 (2017).
12. Qian, X., Nguyen, H. N., Jacob, F., Song, H. & Ming, G. L. *Development* **144**, 952–957 (2017).
13. Quadrato, G. *et al. Nature* **545**, 48–53 (2017).
14. Mansour, A. A. *et al. Nature Biotechnol.* https://doi.org/10.1038/nbt.4127 (2018).
15. Koch, C., Massimini, M., Boly, M. & Tononi, G. *Nature Rev. Neurosci.* **17**, 307–321 (2016).
16. Bourret, R. *et al. Stem Cell. Res. Ther.* **7**, 87 (2016).
17. Greely, H. T., Ramos, K. M. & Grady, C. *Neuron* **92**, 637–641 (2016).
18. *Gray Matters: Integrative Approaches for Neuroscience, Ethics, and Society* (Presidential Commission for the Study of Bioethical Issues, 2014); available at https://go.nature.com/2qvm83f
19. *Gray Matters: Topics at the Intersection of Neuroscience, Ethics, and Society* (Presidential Commission for the Study of Bioethical Issues, 2015); available at https://go.nature.com/2vdqx5j

A full list of author affiliations accompanies this Comment online (see go.nature.com/2hz3ogu).
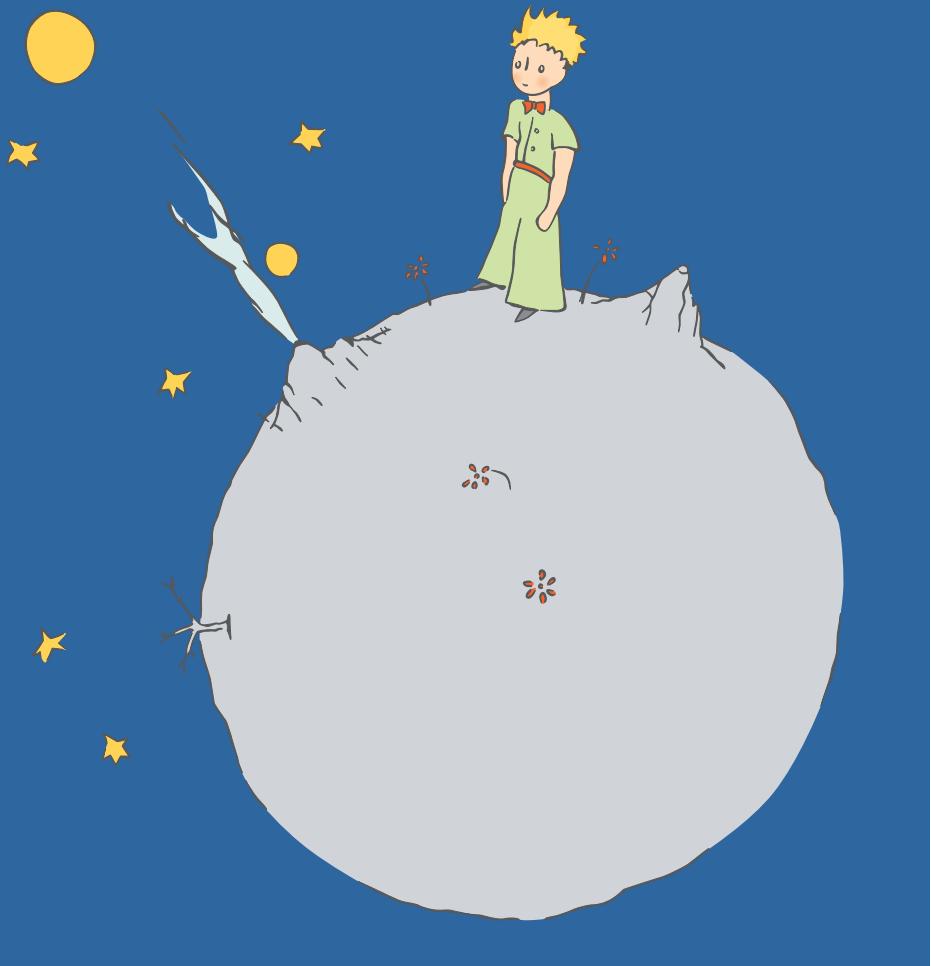
Antoine de Saint-Exupéry's *The Little Prince* teaches the ethics of planetary stewardship.

ENVIRONMENTAL SCIENCE

Stories to save Earth by

S. F. Said explores a study on children's books as preparation for planetary crises.

Today's children will face huge environmental challenges, from climate change to oceanic pollution. The International Union for Conservation of Nature, for instance, has noted that nearly one-quarter of mammals are globally threatened or extinct. In *Beasts at Bedtime*, ecologist Liam Heneghan argues that books can help children deal with these grim eventualities.

Heneghan's assertion is partly a response to the 'No Child Left Inside' movement, sparked by journalist Richard Louv's 2005 *Last Child in the Woods*. Heneghan supports Louv's aim of re-introducing today's digital-drenched children to outdoor life. But he also believes that reading and being read to help children gain environmental literacy, enabling them to engage with

nature in profound ways. To make his case, Heneghan discusses around 20 children's books in detail, and analyses their environmental themes.

His focus is on classics such as Beatrix Potter's *Tale of Peter Rabbit* (1902) and L. Frank Baum's *The Wonderful Wizard of Oz* (1900): British and North American texts that have found a global readership. The selection is drawn from his experiences as a parent and reader, and from recommendations. In lists provided by US professional teachers' organization the National Education Association, for example, he finds that every book recommended for preschoolers is environmentally themed. Meanwhile, 60% of those recommended for 4- to 8-year-olds "feature animals or are in other ways concerned with nature", as do 50% for 9- to 12-year-olds.

Heneghan structures his study according to the settings of the books he chooses: pastoral, wilderness, island and urban. *Beasts at Bedtime* thus feels more like a catalogue than an evolving argument; there is little sense of historical context or development. Heneghan fails to engage, for instance, with why animal stories of the late nineteenth century, such as Anna Sewell's *Black Beauty* (1877) or Rudyard Kipling's The Jungle Books (1894–95), became so popular in a time of urbanization and industrialization. Those trends removed people from daily contact with livestock and wild fauna, exoticizing them.

Heneghan's analyses of individual books can be apt. In Ursula K. Le Guin's Earthsea series — beginning with the 1968 *A Wizard of Earthsea* — he identifies a thread that runs through many fantasy works. He notes that "before extravagant quests, before dragons and gold, before strenuous heroism comes botany. Hobbits farm the Shire, Harry Potter visits the greenhouses with Professor Sprout, and Ged walks the mountains of Gont with Ogion the Silent, learning the uses of plants."

Heneghan shows how Le Guin's concept of magic and wisdom is tied to the capacity to name the natural world, as well as values such as balance, connectedness and responsibility — all fed by her feminism, Taoism and ecological leanings (see M. S. Barr *Nature* **555**, 29; 2018). Although the theory of a "balance of nature" has been rejected by ecologists starting with Aldo Leopold in the 1920s, as Heneghan acknowledges, Le Guin's ideas remain relevant in an era so marked by human disruption of natural systems that some dub it the Anthropocene.

Ultimately, however, the value of this book is limited by its narrow focus on classics. Twenty-first-century children's books are represented only by J. K. Rowling's Harry Potter series and Suzanne Collins's Hunger Games trilogy. Yet hundreds of books over the past decade or so are relevant to his theme, from Michelle Paver's *Wolf Brother* and Piers Torday's *The Last Wild* to M. G. Leonard's *Beetle Boy*.

Heneghan does note that wilderness fiction

Beasts at Bedtime: Revealing the Environmental Wisdom in Children's Literature
LIAM HENEGHAN
University of Chicago Press (2018)

from Daniel Defoe's *Robinson Crusoe* (1719) onwards usually excludes indigenous perspectives, but he does not direct readers to any indigenous children's authors. Louise Erdrich's superb ongoing Birchbark House series, for instance, is a Native American response to Laura Ingalls Wilder's Little House series, fictionalized memoirs published in the early twentieth century.

Erdrich depicts an entire culture training children from toddlerhood to be stewards of nature.

Non-fiction is also regrettably omitted from *Beasts at Bedtime*, especially beautifully illustrated works for children such as last year's *The Lost Words* by Robert Macfarlane and Jackie Morris. This book was sparked in part by a *Science* study finding that children identify Pokémon characters more readily than real flora and fauna (A. Balmford *et al. Science* **295**, 2367; 2002). Macfarlane's poems and Morris's illustrations counter that loss of knowledge by helping children to identify plants and animals, from acorn to wren. Meanwhile, *The Pebble In My Pocket* (1996) and *The Drop In My Drink* (1998) by Meredith Hooper and Chris Coady have made Earth science and the water cycle accessible to young readers. It's hard to imagine projects more relevant to *Beasts at Bedtime*.

These disappointments notwithstanding, Heneghan makes good use of two classics. He uses close readings of Antoine de Saint-Exupéry's *The Little Prince* (1943) and *The Lorax* (1971) by Dr Seuss (pen name of Theodor Geisel) to suggest different approaches to environmentalism for children today.

Heneghan sees *The Little Prince* as "a complete guide to understanding our responsibilities in caring for the world". Saint-Exupéry, an aviator in the Second World War, saw the conventional adult world as built on destructive human folly: illusions of control, narcissism and calculative thinking. Heneghan finds a powerful counterweight in the relationship between the Little Prince and the fox. As the fox tells the prince: "You become responsible, forever, for what you have tamed." Thus, Heneghan identifies an ethic of human obligations towards the planet and its non-human denizens.

By contrast, he reads *The Lorax* as an anti-manifesto: a case study of how environmental advocacy can go wrong, with the Lorax as "a self-righteous, blustering, and ultimately failed environmentalist ... hectoring, stigmatizing and shaming" the environmentally unenlightened Once-ler. The Lorax fails to find commonality with a potential conservation ally and engage them intellectually or emotionally.

That dual engagement is where children's literature can play a decisive part. Story has the power to develop empathy and build knowledge, as well as nurture curiosity and imagination. Childhood reading is undeniably formative, and it's refreshing to see it being taken seriously. Children's books alone cannot save the natural world; but they can spark concern, teach the science and reveal strategies in ways both subtle and direct. ∎

S. F. Said *is a British children's author whose books include* Varjak Paw *and* Phoenix. *Twitter: @whatSFSaid*

Books in brief
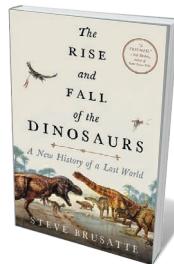
The Rise and Fall of the Dinosaurs
*Steve Brusatte* WILLIAM MORROW *(2018)*
Palaeontologist Steve Brusatte has described more than 15 new species of fossil vertebrate, including the long-snouted theropod "Pinocchio rex" (*Qianzhousaurus sinensis*). In this vivid, pacy chronicle, he meshes the findings in a field currently seeing a new species unearthed, on average, every week with a re-creation of the dinosaurs' 150-million-year reign. This is scientific storytelling at its most visceral, striding with the beasts through their Triassic dawn, Jurassic dominance and abrupt demise in the Cretaceous period, which spared only the theropods from which birds are descended.

Secrets of the Snout: The Dog's Incredible Nose
Frank Rosell (transl. Diane Oatley) UNIVERSITY OF CHICAGO PRESS *(2018)*
With up to 300 million olfactory cells to our 5 million, dogs are spectacularly equipped to sense fugitive compounds emitted by everything from buried mines to colorectal cancer. In this fascinating study, behavioural ecologist Frank Rosell guides us through compelling research on olfaction-related canine ethology, physiology and neuroscience. Interwoven are feats of star sniffer dogs such as Tucker, a seagoing research Labrador that detects killer whales by locating their faeces; and Aska, trained to smell the pheromones of spruce bark beetles, a major insect pest.
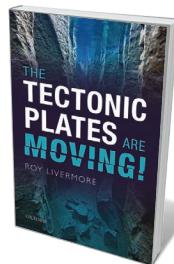
The Tectonic Plates are Moving!
Roy Livermore OXFORD UNIVERSITY PRESS *(2018)*
In 1963, a revolution began to rumble in Earth science. The Vine–Matthews hypothesis (brainchild of marine geologists Frederick Vine and Drummond Matthews) laid the basis for plate tectonics, a key to conundrums as diverse as mountain formation and earthquake location. Consensus came slowly, as geophysicist Roy Livermore charts in this packed account, richly contextualized by the chain of discovery from William Gilbert (author of 1600 treatise *De Magnete*) to Alfred Wegener, Kiyoo Wadati and Ken Bullen. Today's big debates, such as the mechanics of subduction, also get a look-in.
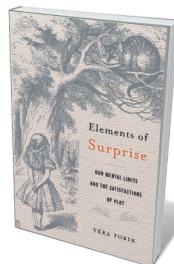
Elements of Surprise
Vera Tobin HARVARD UNIVERSITY PRESS *(2018)*
Plot twists can jolt us into an understanding of fiction's deeper meaning. But how do they work? In this scholarly study, cognitive scientist Vera Tobin pinpoints the psychological quirks that make us vulnerable to literary shock tactics. She shows, for instance, how Charles Dickens harnesses the 'curse of knowledge' bias (the belief that others know what we know) to dizzying effect in his 1861 *Great Expectations*; and how in *Villette* (1853), Charlotte Brontë twists the story twice through unruly protagonist Lucy Snowe, leaving us wallowing in a "vertiginous instability" not unlike Snowe's own.

The Big Cloud
Camille Seaman PRINCETON ARCHITECTURAL PRESS *(2018)*
Photographer Camille Seaman's images of icebergs as entities gnawed by climate change are a window on the world of fast-disappearing polar ice (see J. Hoffman *Nature* **492**, 40; 2012). Here, she turns to a phenomenon even more evanescent: the storm cloud. Carefully avoiding "disaster tourism", Seaman captures stupendous storm fronts, from supercells to baby tornadoes, across South Dakota, Kansas, Nebraska — a record of meteorology under the cosh of a shifting climate, and a homage to untameable nature. **Barbara Kiser**

Correspondence

Order must spring from chaos in Italy

Italy's election last month does not bode well for science there, which is still beset with the problems you highlighted in a feature 35 years ago (*Nature* **303**, 109–128; 1983). It is meagre consolation that research output is increasing, against the odds (*Nature* **554**, 411–412; 2018).

We call on the new political leaders to work with internationally renowned scientists on a major reform of the Italian research system. This needs to be based on merit, transparency and internationalization. Funding should be doubled from its current 1.2% of gross domestic product to be on a par with that of other developed countries.

Four reforms are crucial. Italy needs an independent, public research foundation that supports centres of excellence as part of the European network. This should award funding that is open to all researchers, irrespective of affiliation, and subject to international peer review. Italian universities should be released from public administration to reduce bureaucracy and to allow them autonomy in hiring. Public funds must be disbursed to universities strictly on the basis of the performance of the teaching and research faculty, both to be independently assessed. And the public administration should appoint only PhD-trained skilled professionals to senior posts.

Gerry Melino* *University of Rome Tor Vergata, Italy.*
gm614@mrc-tox.cam.ac.uk
*On behalf of 9 correspondents (see go.nature.com/2jzk41e for full list).

Exercise caution in cod management

Canada's stock of northern cod (*Gadus morhua*) off Newfoundland and Labrador has made a remarkable recovery over the past decade, but remains well below historical levels and current conservation limits. Despite scientific advice to minimize removals (see, for example, S. Rowe and G. A. Rose *Nature* **545**, 412; 2017), the government bowed to political pressure and the reported catch for 2017 was almost triple that for 2015.

Last month's assessment concludes that spawning stock biomass decreased by about 30% over the past year, and predicts a high probability of continued decline for 2019. This stalled productivity, together with evidence that the government's 2017 fisheries-management plan failed to meet the target of sustained stock growth (see go.nature.com/2qt3trk), make it incumbent on Canada's government to reduce cod mortality from fishing.

In our view, maintaining the current harvesting level or continuing expansion will jeopardize long-term stock recovery and a rebuilt fishery. Both stand to deliver another black eye to Canadian fisheries management.

Sherrylynn Rowe *Memorial University of Newfoundland, St. John's, Canada.*
George A. Rose *University of British Columbia, Vancouver, Canada.*
sherrylynn.rowe@mi.mun.ca

Geoengineering is not a quick glacier fix

We disagree with John Moore and colleagues that geoengineering could counter rising sea levels from the melting of the Greenland and Antarctic ice sheets (*Nature* **555**, 303–305; 2018). As environmental researchers in these regions, we contend that the consequences of the technology could be even more serious than in its absence.

The authors' suggestions (building ocean-bottom sills, installing pinning-point islands and removing subglacial water) might briefly slow outflow. However, these strategies could easily cause ice build-up that would overwhelm structural impediments, and further accelerate ice loss.

Even if feasible, slowing the flow of glaciers such as Jakobshavn Isbrae is only a partial solution, given that more than half of ice loss in Greenland is due to surface melt. And the logistical difficulties of transporting the unprecedented amounts of equipment and materials required needs to be taken into account. The Amundsen Sea in western Antarctica, for example, is often inaccessible to icebreakers. And the water-pumping system Moore *et al.* propose would have to extend over almost the entire glacier catchment to avoid water pooling or redirection.

In our view, the limited resources available should instead be used to address the root causes of accelerating ice loss — namely emissions and human-induced climate change.

Twila A. Moon* *National Snow and Ice Data Center, CIRES, University of Colorado, Boulder, USA.*
twila.moon@nsidc.org
*On behalf of 7 correspondents (see go.nature.com/2hhedct for full list).

Code of conduct for research integrity

I welcome the timely release of a Code of Ethics for Researchers by the World Economic Forum (WEF; see *Nature* **555**, 5; 2018). I am concerned, however, that the simplicity of the seven principles could render this code unusable by researchers.

As chair of the drafting group for the European Code of Conduct for Research Integrity, published by ALLEA (All European Academies), I recommend that scientists read the WEF's code of ethics in parallel with other codes, such as ours (see go.nature.com/2hdestq). The European code is built on the views of a range of stakeholder organizations such as the European Network of Research Integrity Offices, ensuring comprehensive advice on how to implement good research practice.

For example, the WEF's code encourages 'engagement with decision-makers'. This might not be easy for researchers who are unfamiliar with policy windows or with the best way to present evidence to decision-makers who are trying to balance many other inputs. The European code offers practical guidance, and the 2017 Brussels Declaration (see go.nature.com/2exd69c) sets out responsibilities for all actors, including researchers, to ensure reliable, evidence-based policymaking.

Maura Hiney *Health Research Board Ireland, Dublin, Ireland.*
mhiney@hrb.ie

Maryland's pursuit of fair drug prices

Bob More calls on the drug industry to hold managers accountable if they engage in opportunistic pricing (*Nature* **555**, 561; 2018). In the meantime, the industry should embrace progressive policy measures that would protect people by demanding transparency from drug firms.

Maryland's General Assembly is leading the way, having passed a landmark law last year to prosecute companies that increase the prices of off-patent drugs. This year, the legislature aims go a step further and create a commission to review drug costs that insists on transparency from pharmaceutical firms to ensure fair pricing for consumers. Companies would be required to report and justify high prices for new drugs and large price increases for older ones (see go.nature.com/2hxujdt).

Rather than supporting such efforts to curb bad practices, however, trade associations such as the US Biotechnology Innovation Organization and PhRMA (Pharmaceutical Research and Manufacturers of America) have sent lobbyists to Maryland's state capital to oppose such measures (see go.nature.com/2vrazr1).

Leah Cairns *Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.*
lcairns3@jhu.edu

ORGANIC CHEMISTRY

Facial recognition for molecules

A catalyst has been developed that recognizes the topology of just one face of a planar reaction intermediate. Remarkably, this enables one mirror–image isomer of the reaction product to be made selectively. See Article p.447

TOBIAS MORACK & RYAN GILMOUR

Complex 3D molecules are ubiquitous in daily life, with functions in everything from high-performance materials to smart medicines. Just as 3D shape often reflects function on the macroscopic scale, so, too, does it determine microscopic behaviour. When building 3D molecules for applications, chemists must therefore develop synthetic routes that ensure that each atom is correctly positioned in the final product. However, even if such precision is achieved, some molecules can be produced as mirror-image isomers (enantiomers), and their properties might vary widely, affecting their use in applications. On page 447, Wendlandt et al.[1] report a reaction that not only enables the synthesis of single enantiomers of molecules, but does so using a reaction mechanism that was thought to be intrinsically lacking in enantioselectivity.

A rich arsenal of synthetic methods is available to connect molecular fragments in a highly predictable manner, and substitution processes are among the most powerful of these. Known as S_N1 or S_N2 reactions, these processes share a common requirement for an electron-rich species (a nucleophile) and an electron-deficient species (an electrophile). However, the two reaction types proceed by distinct mechanisms.

In S_N2 reactions, the approach of the nucleophile towards the reactive centre of the electrophile leads to the extrusion of a chemical group, known as the leaving group (Fig. 1a). The leaving group on the electrophile is then replaced — substituted — by the nucleophile. The bonds around the reaction centre behave much like the spokes of an umbrella turning inside out on a stormy day. This highly orchestrated mechanism ensures that the 3D geometry in the tetrahedral framework of bonds around the reaction centre is inverted, and the 'information' about that geometry, encoded by the framework (the stereochemical information), is not lost. In other words, only one enantiomer of the product is formed.

By contrast, the S_N1 reaction involves the initial extrusion of the leaving group — independently of the nucleophile — to generate a charged species called a carbocation (Fig. 1b).
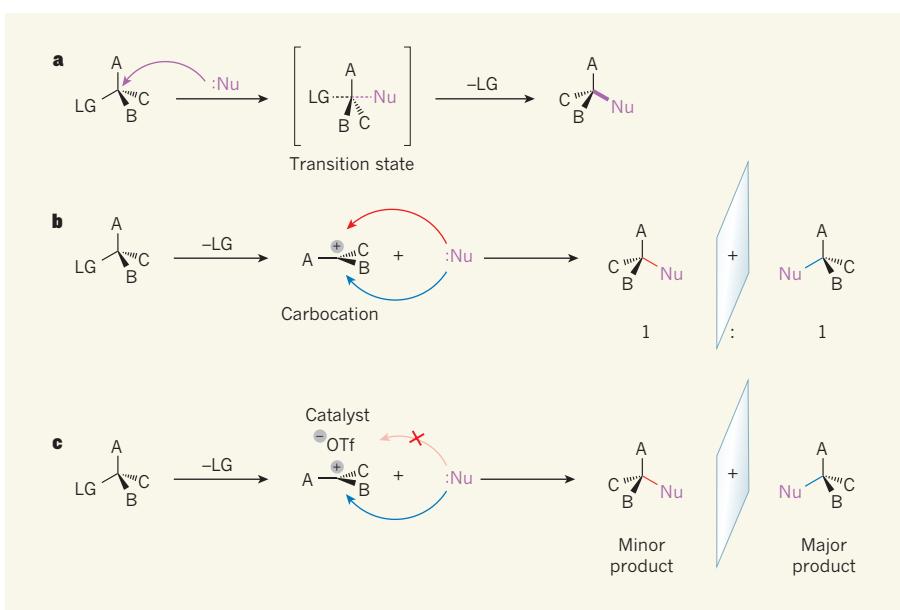

Figure 1 | Mirror-image selectivity in substitution reactions. a, In an S_N2 reaction, a reactant called a nucleophile (Nu; dots indicate a lone pair of electrons) attacks a carbon atom in an organic molecule from the side opposite to a 'leaving' group (LG). A, B and C can be any atom or group. In the transition state, the bond between the carbon atom and the LG is partly broken, and a bond between Nu and the carbon atom is partly formed (partly broken and partly formed bonds are shown as dashed lines). The bond to the LG then breaks, and a single product is formed. **b,** In the S_N1 reaction, the LG is released first, and a planar, charged intermediate called a carbocation forms. Because the Nu can attack the carbocation equally easily from either side, the product forms as a 1:1 ratio of mirror-image isomers (enantiomers). **c,** Wendlandt et al.[1] report an S_N1 reaction in which a small-molecule catalyst and a triflate ion ($^-$OTf) bind to one side of the carbocation, directing the nucleophile to the other side. One enantiomer is therefore produced preferentially.

Because this intermediate is planar, it can be intercepted by the nucleophile from either face, to generate the reaction product as a one-to-one mixture of enantiomers, thus losing the stereochemical information in the electrophile. Wendlandt et al. overcome this long-standing stereochemical limitation by reporting an enantioselective S_N1 reaction.

The authors' method can be thought of as emulating biometric facial recognition: they use a small-molecule catalyst to discriminate between the two faces of the transient carbocation. Much as algorithms efficiently analyse structural features of faces, the catalyst identifies one face of the cation and thus directs the nucleophile to the other (Fig. 1c). This facial discrimination manifests in the high enantioselectivity of the reaction. The formation of one-to-one mixtures of enantiomers is prevented, but, in contrast to the S_N2 reaction,

the stereochemical information in the starting material is irrelevant. The reaction is therefore said to be enantioconvergent.

The small-molecule catalyst used by Wendlandt and colleagues is combined with a Lewis acid (a molecule that accepts electron pairs from other molecules), to generate a negatively charged complex that can be thought of as the active form of the catalyst. This type of combination is common in nature, and often facilitates reactions that would be impossible using the catalyst alone. In the present study, the active catalyst activates electrophiles known as propargyl acetates by removing the leaving group (the acetate), forming a carbocation intermediate. This positively charged intermediate interacts with the negatively charged active catalyst to form an ion pair; such ion pairing has previously been used

in other types of enantioselective catalytic reaction[2]. The active catalyst recognizes one face of the planar cation, shielding it so that only the opposite face can interact with an approaching nucleophile. The products of this S_N1 reaction are thus obtained predominantly as one enantiomer.

Not only does Wendlandt and co-workers' study transform an S_N1 reaction into a catalytic, enantioconvergent process, but it also constitutes a powerful synthetic route for preparing molecular motifs called quaternary carbon centres[3,4], which are notoriously challenging to make enantioselectively. Quaternary carbon centres have four different carbon-based substituents attached to a central carbon atom, and are commonly found in biologically active, naturally occurring compounds, such as morphine or various steroids. A great deal of structural diversity could be generated by varying each of the four substituents, making quaternary carbon centres valuable starting points for drug discovery. The authors' study is a breakthrough in that it allows readily accessible racemic mixtures (one-to-one mixtures of enantiomers) of starting materials to be quickly processed to make structurally complex molecular scaffolds containing these motifs.

Wendlandt et al. exemplify their reaction using substrates that contain structural groups specifically chosen to stabilize the intermediate cation. However, the underlying concept is general, and will certainly be translated to related classes of reaction. The reported products are striking because the carbon atoms attached to the central atom in the quaternary centre represent a diverse range of electron orbitals (sp, sp^2 and sp^3). This means that the groups attached to the central atom have markedly different geometries and reactivities. Future work in which the four groups are varied will therefore produce highly versatile libraries of molecules that could be used in a wide range of reactions for synthesis, and allow the exploration of a large amount of 'chemical space' — the vast array of all possible molecules. ∎

Tobias Morack and **Ryan Gilmour** *are at the Organisch-Chemisches Institut, University of Münster, 48149 Münster, Germany.*
e-mail: ryan.gilmour@uni-muenster.de

1. Wendlandt, A. E., Vangal, P. & Jacobsen, E. N. *Nature* **556**, 447–451 (2018).
2. Brak, K. & Jacobsen, E. N. *Angew. Chem. Int. Edn* **52**, 534–561 (2013).
3. Quasdorf, K. W. & Overman, L. E. *Nature* **516**, 181–191 (2014).
4. Feng, J., Holmes, M. & Krische, M. J. *Chem. Rev.* **117**, 12564–12580 (2017).

EPIDEMIOLOGY

A broader look at paediatric HIV infection

An epidemiological study of adolescents who had acquired HIV around the time of birth highlights how high-income countries benefit from the ability to begin treating all infected children in the first years of life.

PATRICIA M. FLYNN

It is estimated that more than 2.1 million children under 15 years of age are living with HIV infection worldwide (see Statistical Tables at go.nature.com/2qqvksh). Almost all of these children received the retrovirus from their mother, either around the time of birth (the perinatal period) or through breast-feeding. Writing in *PLoS Medicine*, the CIPHER (Collaborative Initiative for Paediatric HIV Education and Research) Global Cohort Collaboration[1] has shed light on the population of children who survive perinatal HIV infection, living into adolescence and beyond. The collaboration's far-reaching epidemiological study emphasizes the importance of early antiretroviral therapy (ART) for improving survival rates and clinical outcomes.

Before the availability of ART, the mortality rate of babies perinatally infected with HIV was 25% by 2 years of age[2]. Today, a much higher percentage of perinatally infected children survive into adolescence, but this population is little studied compared with other cohorts of people with HIV. Furthermore, it has been unclear whether or how perinatally acquired HIV might differentially affect adolescents living in different parts of the world.

The CIPHER investigators combined data from researchers in 12 regions of the world who followed adolescents with perinatal HIV infection. For a total of 38,187 adolescents, the authors report factors such as the age at which ART began, the person's growth record, whether or not they dropped out of their treatment programme, and survival. Individuals were tracked from their first clinic visit and first ART treatment until 15 years of age, and data were collected from the 1980s or 1990s (depending on region) until 2014.

These data allowed the CIPHER researchers to compare adolescents in different countries, regions and country income levels. Their analysis demonstrates that country income level is associated with the stage of life at which ART typically begins. Not surprisingly, the initiation of both care and ART occurred much earlier in the high-income countries included in the study (14 European countries and the United States) than in other areas of the world. Specifically, children first visited a clinic at a median of 1.1 years of age and began ART at a median of 2.5 years in high-income countries, compared with 7.1 years and 8.0 years, respectively, in the low-income countries studied (26 countries in sub-Saharan Africa).

Growth is delayed in children with HIV infection[3], and was markedly different between high- and low-income countries. The authors found that height-for-age scores were higher in high-income countries than in countries at other income levels at the times when the person first visited the clinic and began ART. In all regions, growth levels began to catch up after ART began. But the height-for-age scores achieved by children in high-income countries at the age of 10 and at the time of the last recorded clinic visit were similar to those of children and adolescents who did not have HIV. By contrast, height-for-age scores for
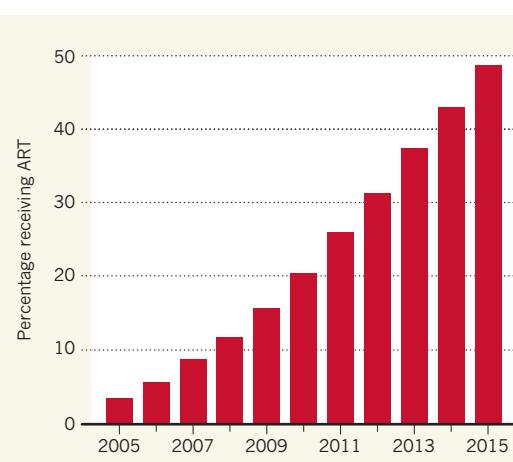

Figure 1 | Prevalence of antiretroviral therapy (ART) for children living with HIV in low- and middle-income countries. Between 2005 and 2015, the percentage of children under 14 years of age who were living with HIV in low- and middle-income countries and who were receiving ART rose from less than 5% to 49%. But the CIPHER Global Cohort Collaboration[1] reports that these children typically began ART later in life than their counterparts in high-income countries.

SOURCE: UNICEF

children living in regions with lower incomes remained below average, probably because ART began significantly later in life. In addition, the cumulative incidence of mortality by age 15 was approximately 3 times greater in low-income countries than in high-income ones (2.6% and 0.9%, respectively).

This data set will no doubt improve our understanding of the population of young people living with HIV infection, in no small part owing to its impressive size and breadth. Previously, the only available data sets included many fewer children and adolescents, and were limited to regional or national cohorts. However, the work also comes with caveats. For instance, the cohorts studied in different parts of the world were established for different reasons — some were national registries, whereas others were groups followed as part of other research. As such, there could be some biases in the study. In addition, there were some differences in the data elements that were collected in different regions. Future studies looking at specific aspects of the data might therefore be limited to fewer adolescents.

The CIPHER investigators' work adds to previous evidence[4,5] for the benefits of early ART. Together, these key studies demonstrate that early ART can save lives and preserve normal growth and development in children with HIV infection. However, the scale-up of ART for infected children has lagged woefully behind that for adults. Data from the United Nations' children's agency Unicef (see go.nature.com/2qqc5gt) show improvements in the availability of ART for children under 15 years of age since 2005, but there is still much to be done. In 2015, fewer than half of infected children were receiving ART (Fig. 1).

There are many barriers to the optimal provision of ART for children in low-income countries[6]. For instance, problems with infrastructure, including a lack of medical personnel and insufficient drug stocks, are common in many areas of the world. Drug development is challenging, because the physiological changes that occur during childhood can affect drug absorption, distribution, metabolism and excretion, and much work is therefore typically required to make drugs safe and effective for children of all ages. ART consists of three drugs that, in adults, can be given in one pill — but this is trickier in children, because the appropriate ratios of the drugs might vary with age. Finally, infant- and child-friendly formulations that do not require refrigeration and are easy to transport are currently lacking.

> "Early antiretroviral therapy can save lives and preserve normal growth and development in children with HIV infection."

Because advances in preventing perinatal HIV infection have markedly decreased the numbers of infected children, there is unfortunately no longer a market benefit to developing paediatric ART, making it an unattractive focus for pharmaceutical companies. The current paper's clear demonstration of the benefits of early ART demands that efforts to develop optimal drugs and formulations be increased. In addition, infrastructure must be put in place to make ART consistently available to the more than one million children with HIV infection who do not currently have adequate access to these live-saving drugs. ∎

Patricia M. Flynn *is in the Department of Infectious Diseases, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA.*
e-mail: pat.flynn@stjude.org

1. The Collaborative Initiative for Paediatric HIV Education and Research (CIPHER) Global Cohort Collaboration et al. *PLoS Med.* **15**, e1002514 (2018).
2. Scott, G. B. *et al. N. Engl. J. Med.* **321**, 1791–1796 (1989).
3. Barlow-Mosha, L. *et al. J. Int. AIDS Soc.* **20**, 21552 (2017).
4. Violari, A. *et al. N. Engl. J. Med.* **359**, 2233–2244 (2008).
5. Laughton, B. *et al. AIDS* **26**, 1685–1690 (2012).
6. Penazzato, M. *et al. Clin. Infect. Dis.* **64**, 1597–1603 (2017).

Shrimp cause a stir

Brine shrimp (*Artemia salina*) are a type of tiny crustacean that lives in swarms and has a daily pattern of vertical migration. In a paper online in *Nature*, Houghton *et al.* report that such group migration generates water eddies with the potential to cause substantial mixing of the water column (I. A. Houghton *et al. Nature* https://doi.org/10.1038/s41586-018-0044-z; 2018).

The ability of individuals or groups to alter their physical environment has long fascinated biologists. Indeed, Charles Darwin's final book, *The Formation of Vegetable Mould through the Action of Worms* (Murray, 1881), reported his analysis of the changes that could occur through the repeated actions of small creatures. This work was a fitting finale for a career spent showing how small changes could, given the time and opportunity, have large effects. As with worms, so, too, with shrimp.

In laboratory-conducted experiments, Houghton and colleagues studied the effect of *A. salina* group migration (pictured: a time-lapse image in which the vertical tracks made by suspended particles provide a way of monitoring water flow). They found that shrimp movement created a water jet that caused mixing of the water column on a scale three orders of magnitude more effective than the mixing that occurs by diffusion. Ocean mixing can be caused by wind or currents. Small marine organisms might also contribute to perceptible ocean-mixing changes, if the phenomenon reported by Houghton and colleagues is relevant for tiny crustaceans called krill, which swarm in vast numbers in climatically sensitive parts of the world, such as the Southern Ocean. **Henry Gee**
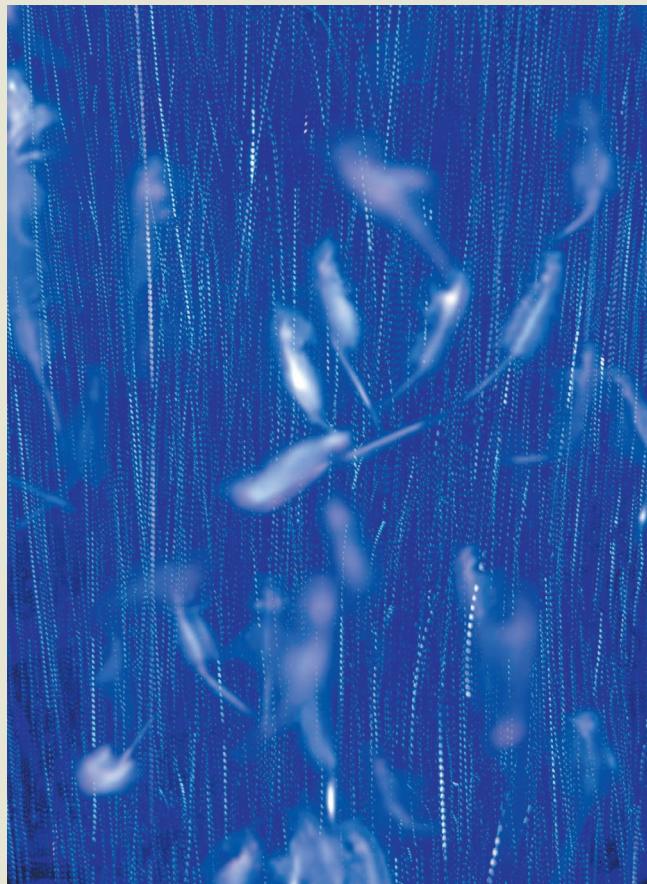
ISABEL HOUGHTON

BIOTECHNOLOGY

Organoids reveal cancer dynamics

Single–cell analyses in cancer are limited by the small biomass of individual cells. *In vitro* **production of 3D organoid structures from single tumour–derived cells generates sufficient biomass for in–depth analyses.** SEE ARTICLE P.457

CALVIN J. KUO & CHRISTINA CURTIS

Cancer is thought to arise from a single cell that proliferates to form a clonal population. Cells in that population then progressively acquire distinct mutations, leading to rampant tumour-cell diversification. Ideally, this intratumoral heterogeneity would be studied at the single-cell level — but such explorations are limited because single cells contain very little material for analysis. On page 457, Roerink *et al.*[1] overcome this limitation using methods for growing 3D cultures from single cells isolated from colorectal cancers. These organoid cultures can be used as proxies for the single cells from which they are derived, enabling in-depth analysis of tumour diversity.

Different studies of intratumoral heterogeneity have inferred distinct models of tumour evolution. According to the 'big bang' model, most detectable heterogeneity arises early in a cancer's evolution, coincident with the cells becoming cancerous. This is followed by proliferation of a range of subclones that are effectively equally fit[2,3]. In another model, tumours diversify gradually, through ongoing sequential selection of subclones that gain mutations conferring increased fitness[3–5]. Regardless of how it arises, heterogeneity is inextricably linked with intratumoral differences in growth, aggressiveness and sensitivity to therapy, so an in-depth understanding of the phenomenon would be valuable[6,7].

Organoid techniques were initially developed to grow normal tissues *in vitro* in 3D, but have proved equally useful for propagating clinical tumour specimens[8–12]. Roerink *et al.* set out to examine intratumoral heterogeneity using organoids. The authors took tissue from between four and six sites in tumours from three patients, and isolated single cells from each site (Fig. 1). Growth of these cells in specific culture conditions produced clonal organoids, in which all cells are nominally identical. In addition, the researchers grew clonal organoids from cells taken from the normal colorectal tissue surrounding each tumour, for comparison. The essentially unlimited proliferation of these single-cell-derived organoids generated sufficient material for deep genomic analysis, yielding numerous insightful conclusions.

The authors analysed the mutations in multiple cancer- and healthy-tissue-derived clonal organoids, and used this information to construct phylogenetic trees describing the lineage relationships between individual cells in a given tumour. In general, the trees recapitulated the geographic diversity of the sites from which the cells were taken. However, Roerink *et al.* found that clonal organoids from the same site were highly diverse, reflecting heterogeneity at the single-cell level.

The authors found mutations common in colorectal cancer (such as those in the genes *APC*, *KRAS* and *TP53*) in the trunks of the phylogenetic trees — that is, present in all cancer-derived organoids from a given person. But most mutations resided in the distal branches of the phylogenetic tree, arising after the cancer had undergone rapid proliferation. Overall, these results might be seen as providing support for the big-bang model in colorectal cancer, in which potential aggressiveness is determined by the early accrual of crucial driver mutations, rather than by the mutations that accumulate later in tumour growth[2].

Next, Roerink *et al.* evaluated the mutational processes associated with tumour progression by analysing mutational signatures and then estimating the contribution of different signatures (which indicate different mutational processes) to the trunk and branches of phylogenetic trees. Certain mutational processes were acting in cancer cells but largely absent in healthy colon cells — and many more mutations were present in the cancer cells. Mutational signatures also differed between the branches of an individual tumour, highlighting the contribution of different mutational processes across the tumour. On the basis of the mutational signatures and mutation burden, the authors inferred that the rates of both cell division and mutations per cell division were elevated in cancers compared with healthy cells. Consistent with an elevated mutation rate, the phylogenetic trees reveal rapid diversification in the tumour genome after the first cell becomes cancerous.

In addition to genomic data from each organoid, the authors profiled RNA transcripts and patterns of DNA modification by methyl groups (which can lead to altered gene expression). The way in which these profiles diversified across organoids closely traced the mutation-based phylogenies for each tumour. One key concern about organoid systems is that they cut cells off from the cues normally provided by the surrounding tissue — this might alter gene-expression or methylation patterns. That Roerink *et al.* observe stable diversification of these patterns indicates that tumour cells in organoids no longer rely on the surrounding microenvironment to maintain gene expression or methylation, highlighting the power of this experimental system.

The researchers' organoids also allowed them to analyse how mutational landscapes affect drug resistance. Certain targeted
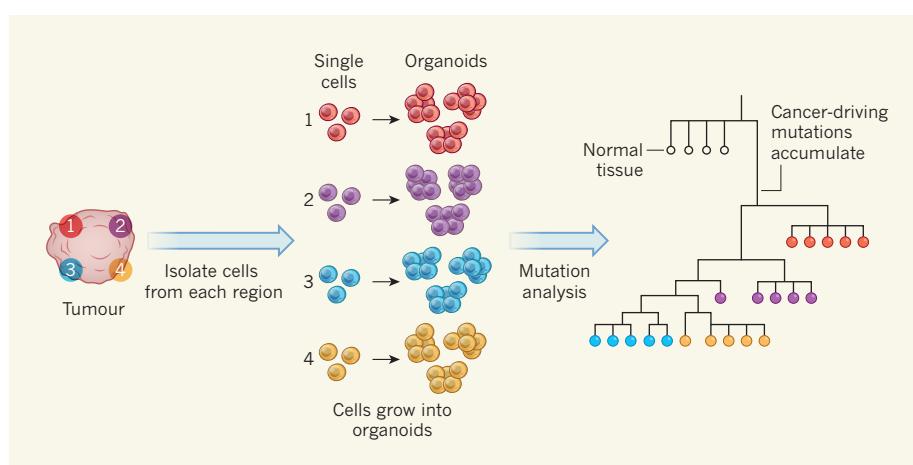

Figure 1 | Cancer analyses using organoids. Roerink *et al.*[1] extracted tissue from several sites from human colorectal cancers. They induced single cells isolated from these samples, and from nearby healthy tissue (not shown), to grow *in vitro* to form 3D cultures called organoids. They then performed a range of analyses, including DNA sequencing, on each organoid. The authors used the mutations in each organoid to construct phylogenetic trees that showed lineage relationships between the organoids' genomes. Those derived from the same site (indicated by colour coding) bore most similarities. The analysis revealed that typical cancer-driving mutations arose in the trunk of the tree (that is, during early stages of tumour growth), but that diversification continued throughout tumour growth. (Tree adapted from Fig. 1 of ref. 1.)

therapies acted as might be expected — for instance, nutlin-3a, which inhibits an interaction between wild-type p53 protein (which is encoded by *TP53*) and a binding partner, effectively killed cells in organoids in which *TP53* was not mutated. By contrast, sensitivity to general chemotherapies did not correlate well with mutational status. The heterogeneity of drug response among samples from different regions of the same tumour provides a particularly sobering demonstration of the challenges of therapeutic resistance in cancer.

Overall, Roerink and colleagues' study is a remarkable demonstration of the ability of organoid technologies to amplify rare cancer-cell populations and enable deep analyses. However, the current approaches have limitations. For instance, organoids do not provide a wholly accurate snapshot of the situation in a tumour, because the growth of organoids involves cell proliferation, which itself can lead to further mutations. To get around this problem, Roerink *et al.* focused on

mutations present in all cells of an organoid in their analysis. Another caveat is that organoid culture methods might also select for cells particularly adapted for *in vitro* growth, potentially correlating with the most proliferative or aggressive cells.

Nevertheless, this and other organoid methods should in the future prove invaluable tools for the genetic and functional dissection of cancer. A big help in this endeavour will be biobanks of patient-derived organoids[9,10,13], which have been established in the past few years to encapsulate the genetic diversity of cancers from different anatomical sites. Eventually, organoids might also incorporate components of the tumour microenvironment, enabling more-holistic cancer modelling. Finally, Roerink and colleagues' method could be generalized beyond cancer to enable the expansion of limiting amounts of biopsy tissue from inflammatory, metabolic or genetic disorders. Here, as in cancer, this approach could reveal cellular heterogeneity and inform treatment strategies. ■

Calvin J. Kuo *is in the Division of Hematology, Stanford University School of Medicine, Stanford Cancer Institute, Stanford, California 94305, USA.* **Christina Curtis** *is in the Division of Oncology and the Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA.*
e-mails: cjkuo@stanford.edu;
cncurtis@stanford.edu

1. Roerink, S. F. *et al. Nature* **556**, 457–462 (2018).
2. Sottoriva, A. *et al. Nature Genet.* **47**, 209–216 (2015).
3. Sun, R. *et al. Nature Genet.* **49**, 1015–1024 (2017).
4. Vogelstein, B. *et al. Science* **339**, 1546–1558 (2013).
5. Nik-Zainal, S. *et al. Cell* **149**, 994–1007 (2012).
6. Bivona, T. G. & Doebele, R. C. *Nature Med.* **22**, 472–478 (2016).
7. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. *Cell* **168**, 707–723 (2017).
8. Gao, D. *et al. Cell* **159**, 176–187 (2014).
9. van de Wetering, M. *et al. Cell* **161**, 933–945 (2015).
10. Boj, S. F. *et al. Cell* **160**, 324–338 (2015).
11. Fujii, M. *et al. Cell Stem Cell* **18**, 827–838 (2016).
12. Sachs, N. *et al. Cell* **172**, 373–386 (2018).
13. Sato, T. *et al. Gastroenterology* **141**, 1762–1772 (2011).

TUMOUR BIOLOGY

Transition states that allow cancer to spread

Cancers of epithelial–cell origin often contain some tumour cells that have acquired traits of mesenchymal cells. How this leads to cancer spread has now been illuminated in mouse models. SEE ARTICLE P.463

**ERIK W. THOMPSON
& SHIVASHANKAR H. NAGARAJ**

In cancers that arise from epithelial cells, some of the tumour cells adopt features of mesenchymal cells and lose their epithelial characteristics. This phenomenon is known as the epithelial-to-mesenchymal transition (EMT), and the emergence of cells with mesenchymal features is often associated with a poorer prognosis for patients. How this transition occurs and the implications of EMT for metastasis, the process by which cancer spreads, are not fully understood. On page 463 , Pastushenko *et al.*[1] report an analysis of mouse tumours undergoing EMT, revealing that distinct cell populations can be identified during this transition.

Cancers of epithelial tissues are called carcinomas and can give rise to cells with mesenchymal properties, such as an elongated shape. These mesenchymal cells have been implicated[2] in tumour metastasis. The EMT process is also associated with cancer stem cells found in breast cancer and other cancer types, and is linked to the presence of the circulating tumour cells in the bloodstream that are a

hallmark of tumour-relapse risk[3]. In numerous mouse models[2], prevention of EMT or its reversal — a process known as MET, in which mesenchymal cells transition back to epithelial cells — can reduce the stem-cell-like properties of tumour cells and thereby reduce their ability to metastasize.

Although some studies[4,5] indicate that not all tumour metastases require EMT, considerable amounts of data implicate[6] the process of MET as a requirement for metastatic cells to successfully colonize locations beyond the initial tumour site. One way to address this controversy would be to gain a better understanding of how the EMT process occurs.

A hybrid EMT state, also known as a metastable state or partial EMT, in which individual cells express both epithelial and mesenchymal markers, has often been observed in carcinoma[7], developmental processes and wound healing[8]. One model for how EMT might occur is that cells make a gradual transition from one state to the other along a continuous spectrum of change in which cells lose epithelial characteristics and concurrently gain mesenchymal ones. This is consistent with gene-expression analysis of cell lines

from various types of carcinoma[9]. However, computational-modelling studies[7] and *in vitro* observations[8] indicate instead that distinctive, stable, long-lasting cell populations could represent discrete intermediate stages of EMT.

Pastushenko and colleagues investigated whether stable EMT states in tumours could be identified *in vivo*. The authors used genetically engineered mice that provide a model system in which to study skin carcinoma and have tumour cells[10] that are known to exhibit EMT. Pastushenko *et al.* studied cells that had lost expression of the epithelial-cell marker EpCAM, and analysed the expression of more than 170 cell-surface proteins to try to identify useful markers of cell subpopulations. They then focused on three of these markers (CD61, CD51 and CD106) that are characteristic of a mesenchymal-cell state. This enabled the authors to identify distinct cell populations that had lost expression of the epithelial-cell marker EpCAM and expressed different combinations of these three mesenchymal markers.

Using the three markers as a tool to enable the isolation of cell populations present when EMT occurs, the authors isolated and characterized six distinctive cell subpopulations that lacked EpCAM expression. The authors used an impressive array of biomolecular analyses to assess the cellular characteristics of these populations, including their metastatic traits and their state of cellular differentiation. This work provides *in vivo* evidence of stable cell populations representing intermediate EMT states.

The authors also identified comparable cell subpopulations, characterized by the same markers, in both a mouse mammary-tumour model and human tumour samples transplanted into mice, suggesting that these subpopulations represent stable transitional cellular states found in different cancer types.
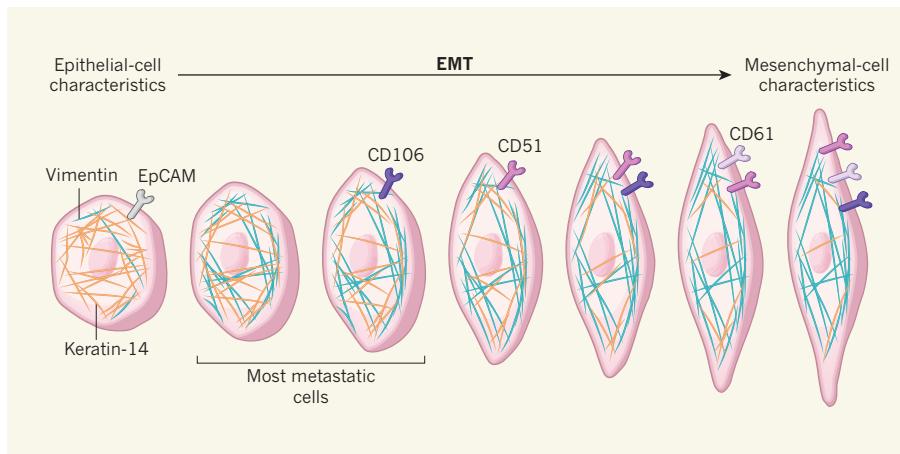
Figure 1 | Populations of cancer cells in transition states. Cancers that arise from epithelial cells often contain some tumour cells that have acquired mesenchymal-cell characteristics, such as an elongated shape and the expression of genes associated with mesenchymal cells. This change is termed the epithelial-to-mesenchymal transition (EMT). Whether this transition occurs as a gradual, continuous change or by discrete stages is debated. Pastushenko *et al.*[1] report mouse studies consistent with the latter, in which they identified cell populations *in vivo* that represent stable, distinct intermediates in this transition. The authors analysed cells that do not express the epithelial-cell receptor protein EpCAM, and isolated cell populations on the basis of their expression of the mesenchymal-cell receptor proteins CD51, CD61 and CD106. They monitored the level of proteins characteristic of the epithelial-cell state (such as keratin-14, yellow) or of the mesenchymal-cell state (such as vimentin, blue), and found cells that exhibited a hybrid of epithelial and mesenchymal characteristics. Cells that expressed EpCAM had low levels of vimentin and high levels of keratin-14. Once EpCAM expression was lost, the level of vimentin increased sharply and the level of keratin-14 declined incrementally as cell populations became more mesenchymal in character. The authors identified two cell populations that were the most likely to spread and form tumours elsewhere in the process called metastasis.

This finding might have clinical relevance. The authors' work therefore provides a platform for future investigations of this phenomenon in other cancer models, including studies of the regulatory networks that control these states. Such investigations might improve understanding of the molecular and cellular features that underlie the tumour-cell capabilities associated with these states.

Do these subpopulations exhibit behaviours that might offer therapeutic opportunities? One of the authors' most striking findings, made in the mouse skin-carcinoma and mammary-tumour models, is that when comparing these cell subpopulations, two of them stood out as having a substantially higher metastatic potential (Fig. 1). This raises the question of whether these highly metastatic cells could be specifically targeted with drugs to arrest tumour progression. Interestingly, none of the stable subpopulations was superior to the others in terms of tumour-initiating capability or proliferation rate. Also, each of the subpopulations that had acquired mesenchymal characteristics exceeded the metastatic capabilities of the epithelial-cell population. These findings are relevant to current debates in this field, which include the question of whether a cycle of EMT followed by MET is required for successful metastatic colonization of a secondary location[6,11].

The authors' work reveals the progressive acquisition of EMT features in EMT hybrid cells, supporting the previously proposed[8] idea of stable transitional states. A process of stepwise change is supported by Pastushenko and colleagues' data, indicating that distinct transcriptional and signalling processes govern intermediate aspects of EMT. The different cell subpopulations had characteristic transcriptional signatures regulated by distinct transcription-factor proteins, underscoring the reproducible and meticulously regulated nature of these hybrid cells.

Although these hybrid cell populations are a stable presence in tumours, the authors found that these cells also maintain a high degree of plasticity, given their ability to undergo MET and revert back to an epithelial-cell state. However, the authors found that the cell subpopulations that were best at undergoing MET during tumour metastasis to the lung in mice were not the most metastatic populations, which will fuel the debate[6] about whether MET is a requirement for metastasis.

Elucidation of the circuitry and signalling feedback loops that might stabilize[7,12] these distinct cellular states will be a worthy goal for future work. Studies of such cellular intermediates in patients' tumours, circulating tumour cells and metastases would be an ideal way to extend this work in a clinical context. If EMT states can be characterized for a wide range of human tumours, this might offer a way to enhance personalized approaches for cancer treatment. The development of specific technologies to identify human EMT states might allow clinicians to predict a tumour's metastatic potential and thereby plan the most effective treatment regimen. ■

Erik W. Thompson *and*
Shivashankar H. Nagaraj *are in the Institute
of Health and Biomedical Innovation, and the
School of Biomedical Sciences, Queensland
University of Technology, Brisbane QLD 4059,
Australia. They are also at the Translational
Research Institute, Woolloongabba, Australia.*
e-mails: e2.thompson@qut.edu.au;
shiv.nagaraj@qut.edu.au

1. Pastushenko, I. *et al. Nature* **556**, 463–468 (2018).
2. Chaffer, C. L., San Juan, B. P., Lim, E. & Weinberg, R. A. *Cancer Metastasis Rev.* **35**, 645–654 (2016).
3. Francart, M. E. *et al. Dev. Dyn.* **247**, 432–450 (2017).
4. Zheng, X. *et al. Nature* **527**, 525–530 (2015).
5. Fischer, K. R. *et al. Nature* **527**, 472–476 (2015).
6. Brabletz, T., Kalluri, R., Nieto, M. A. & Weinberg, R. A. *Nature Rev. Cancer* **18**, 128–134 (2018).
7. Jolly, M. K. *et al. Oncotarget* **7**, 27067–27084 (2016).
8. Nieto, M. A., Huang, R. Y.-J., Jackson, R. A. & Thiery, J. P. *Cell* **166**, 21–45 (2016).
9. Tan, T. Z. *et al. EMBO Mol. Med.* **6**, 1279–1293 (2014).
10. Latil, M. *et al. Cell Stem Cell* **20**, 191–204 (2017).
11. van Denderen, B. J. W. & Thompson, E. W. *Nature* **493**, 487–488 (2013).
12. Bracken, C. P. *et al. Cancer Res.* **68**, 7846–7854 (2008).

This article was published online on 18 April 2018.

QUANTUM PHYSICS

Entangled vibrations in mechanical oscillators

**Two experiments have demonstrated entanglement — non-classical correlations —
between remote mechanical systems comprising billions of atoms. The results
could advance our understanding of quantum physics. SEE LETTERS P.473 & P.478**

ANDREW ARMOUR

In the quantum world, the properties of particles can be correlated in an extremely strange way. Measurements on one particle can influence the properties of another, even if the two particles are far apart. Such behaviour is known as entanglement and was initially so paradoxical that many people, including Albert Einstein[1], thought that the underlying theory must be incomplete. However, experiments have verified the counter-intuitive properties of entanglement[2], and physicists have got used to the idea, recognizing that it could be exploited to develop innovative forms of technology. On pages 473 and 478, respectively,

Riedinger *et al.*[3] and Ockeloen–Korppi *et al.*[4] take the exploration of entanglement in a new direction. They entangle the vibrations of a pair of remote mechanical oscillators, each of which contains billions of atoms.

Mechanical oscillators, such as a mass on a spring or the head of a drum, are familiar objects. They respond to being pushed out of equilibrium by vibrating back and forth at a fixed frequency. Because motion can be generated in lots of different ways — using light, electrical currents or even gravity — mechanical oscillators are highly versatile. Consequently, they have many applications, for example in the detection of weak forces[5].

Uncovering signatures of quantum
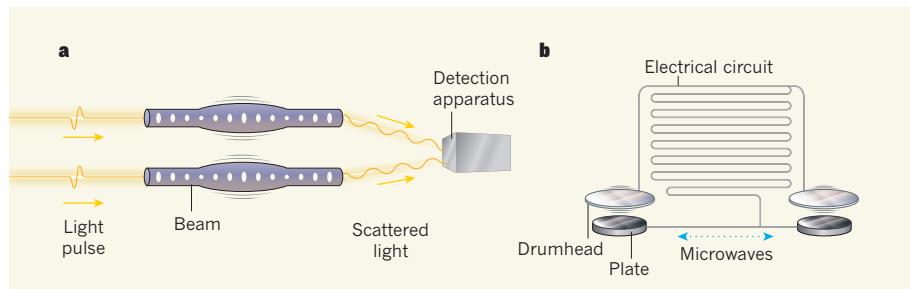
Figure 1 | Entanglement of two types of mechanical oscillator. Riedinger *et al.*[3] and Ockeloen–Korppi *et al.*[4] report entanglement (non-classical correlations) between the vibrations of two remote micrometre-scale mechanical oscillators. **a**, Riedinger and colleagues used oscillators in the form of silicon beams. Each beam contained small holes designed to trap light, that coupled to rapid oscillations in the beam's width. The authors achieved entanglement by shining pulses of light on the beams and detecting the light that was scattered. **b**, By contrast, Ockeloen–Korppi and colleagues used metal drumheads that vibrated up and down above fixed metal plates. The drumheads and the plates were connected by an electrical circuit. The authors injected microwaves into the circuit; these bounced back and forth between the oscillators, coupling the drumheads and giving rise to entanglement.

behaviour, such as entanglement, in mechanical oscillators is exceptionally challenging — largely because it is difficult to prevent these objects from being disturbed by their surroundings. The oscillators usually have low vibrational frequencies and are therefore susceptible to disruption from the thermal jiggling of surrounding atoms. By contrast, the electromagnetic field associated with light has extremely high-frequency oscillations, which means that light is completely insensitive to thermal fluctuations at room temperature. As a result, it is relatively easy to control the properties of light with the exquisite precision required to reveal quantum effects, and the production of entangled light has become almost routine.

The experiments of Riedinger *et al.* and Ockeloen–Korppi *et al.* differed in detail, but shared several key ingredients. To counter the effects of thermal fluctuations, both groups of authors used micrometre-scale mechanical oscillators, which ensured that the vibrational frequencies were not too low, and cooled the oscillators to temperatures of less than 0.1 kelvin. In both experiments, electromagnetic radiation (in the form of light or microwaves) provided the means to generate and detect the entanglement of the oscillators[6,7].

Riedinger and colleagues used a pair of oscillators in the form of 10-μm-long silicon beams — rods that were clamped at both ends and suspended in the middle (Fig. 1a). Each beam contained small holes designed to trap light, that coupled to rapid oscillations (with frequencies of about 5 gigahertz) in the beam's width. The authors shone weak pulses of light on the beams, and monitored the light that was scattered, using a sophisticated scheme that did not reveal which beam the light came from. The detection of such light meant that energy had been transferred from a pulse to the vibrations of a beam, but because there was no information about which oscillator was involved, the vibrations of the two beams were entangled.

The trick of using light to generate entanglement in this way[8] works only if the light scatters from objects that are almost perfect copies of each other. This is difficult to achieve using small mechanical beams, because such objects are produced by a destructive process in which they are essentially sculpted out of a monolithic slab of material. Riedinger *et al.* therefore produced chips containing hundreds of beams from which they selected the best-matched pair.

Ockeloen–Korppi *et al.* used a pair of metal drumheads that vibrated up and down above fixed metal plates (Fig. 1b). The drumheads had diameters of about 15 μm and low vibrational frequencies (about 10 MHz). The authors connected the drumheads by an electrical circuit in which microwaves could bounce back and forth. The microwaves influenced the motion of the drumheads, but were

also affected by this motion, coupling the oscillators in the same way that a spring can link two pendulums. This allowed an entangled state to form, and to persist indefinitely, despite the low vibrational frequencies of the drumheads[9].

Taken together, these two experiments provide an elegant illustration of the power and versatility of electromagnetic radiation as a tool for exploring quantum features of mechanical motion. Each experiment has its advantages. Riedinger and colleagues' beams interface directly with light and are not connected by wires, which means that these devices could be readily integrated into future optical communication networks designed to exploit the effects of entanglement. Ockeloen–Korppi and colleagues' results are particularly striking, given the low vibrational frequencies that they used; and their approach avoids the need for mass fabrication, because the oscillators need not be almost identical.

It was only in 2009 that entanglement was first reported between mechanical oscillators consisting of just two atomic ions[10]. Since then, experiments have demonstrated entangled vibrations in the lattices of crystals[11], at frequencies much higher than even those of Riedinger and colleagues' beams. In terms of the number of atoms involved, the oscillators used by Riedinger *et al.* and Ockeloen–Korppi *et al.* are both a big step up from atomic

ions, but they are still much smaller than the macroscopic objects encountered in everyday life. It will be fascinating to see how much further up in scale experiments are able to go in the next decade. Such progress could lead to exciting insights — for example, larger mechanical oscillators in entangled states might provide answers to outstanding questions about how gravity relates to quantum physics[12]. ∎

Andrew Armour *is in the School of Physics and Astronomy, and the Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham NG7 2RD, UK.*
e-mail: andrew.armour@nottingham.ac.uk

1. Einstein, A., Podolsky, B. & Rosen, N. *Phys. Rev.* **47**, 777–780 (1935).
2. Aspect, A., Dalibard, J. & Roger, G. *Phys. Rev. Lett.* **49**, 1804–1807 (1982).
3. Riedinger, R. *et al. Nature* **556**, 473–477 (2018).
4. Ockeloen–Korppi, C. F. *et al. Nature* **556**, 478–482 (2018).
5. Moser, J. *et al. Nature Nanotechnol.* **8**, 493–496 (2013).
6. Mancini, S., Giovannetti, V., Vitali, D. & Tombesi, P. *Phys. Rev. Lett.* **88**, 120401 (2002).
7. Aspelmeyer, M., Kippenberg, T. J. & Marquardt, F. *Rev. Mod. Phys.* **86**, 1391–1452 (2014).
8. Duan, L.-M., Lukin, M. D., Cirac, J. I. & Zoller, P. *Nature* **414**, 413–418 (2001).
9. Woolley, M. J. & Clerk, A. A. *Phys. Rev. A* **89**, 063805 (2014).
10. Jost, J. D. *et al. Nature* **459**, 683–685 (2009).
11. Lee, K. C. *et al. Science* **334**, 1253–1256 (2011).
12. Blencowe, M. *Nature* **471**, 168–169 (2011).

NEUROSCIENCE

Hunger is a gatekeeper of pain

A neuronal population has now been found that regulates two competing needs — hunger and pain. Urgent pain overrides hunger, but appetite-inducing neuronal activity dampens long-term pain responses to enable feeding.

ALEXEY PONOMARENKO & TATIANA KOROTKOVA

The body's basic needs include a timely supply of nutrients and the avoidance of tissue damage, which are signalled in the brain by hunger and pain, respectively. But these needs cannot be fulfilled simultaneously, because their resolution involves mutually exclusive behaviours. How does the brain prioritize the more urgent need? Writing in *Cell*, Alhadeff *et al.*[1] report that the brain's priorities are set depending on the type of pain involved. Hunger-mediating neurons suppress long-term inflammatory pain, but acute pain, which signals an immediate threat, dampens the activity of these neurons and thus deprioritizes feeding.

Alhadeff and colleagues deprived mice

of food for 24 hours, and analysed how the hungry animals responded to pain. The researchers found that responses to long-term inflammatory pain — of the type associated with chronic disease and recovery from injury — were reduced in the food-deprived animals compared with controls. By contrast, short-term responses to acute pain that was induced by chemicals, heat or force remained intact in hungry mice.

The brain's hypothalamus contains several structures involved in regulating food intake. One of these, the arcuate nucleus, harbours a population of neurons that express agouti-related protein (AgRP), and help to signal nutritional needs — activation of these neurons evokes voracious feeding[2], whereas their ablation leads to starvation[3,4]. Alhadeff *et al.* found that stimulation of the AgRP-expressing
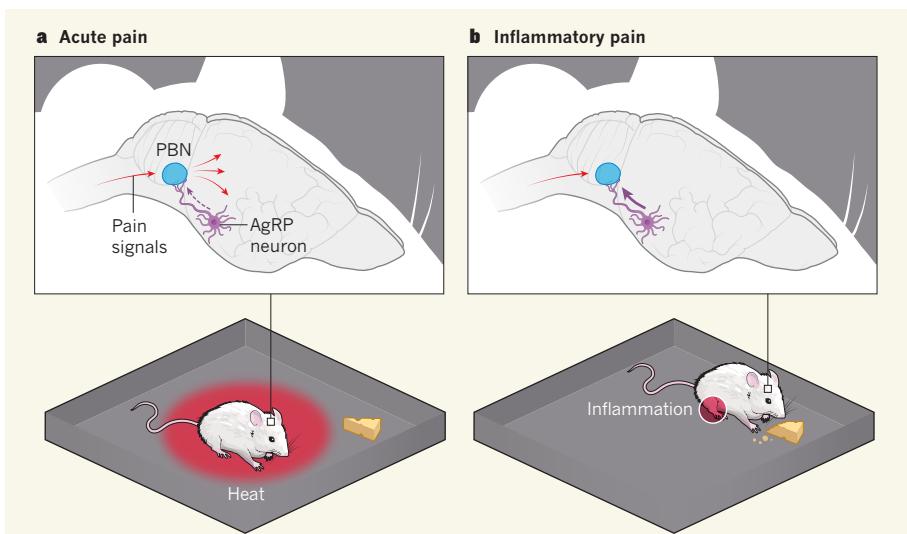
Figure 1 | Getting priorities right in the brain. Alhadeff *et al.*[1] have described a population of neurons that express agouti-related protein (AgRP) and regulate the competing needs of hunger and pain in the mouse brain. **a**, When a mouse is subject to acute pain, AgRP-expressing neurons are inhibited (dashed arrow), and feeding is suppressed. Pain signals from the spinal cord are transmitted throughout the brain via a region called the parabrachial nucleus (PBN). **b**, By contrast, AgRP-expressing neurons remain active during long-term pain, such as that caused by inflammation. The neurons send signals to the PBN to prevent pain transmission to other brain regions, and so feeding is supported.

neurons mimicked the pain-inhibiting effect of hunger in mice. By contrast, silencing of these cells blocked the reduction of inflammatory pain.

AgRP cells send projections to many brain regions. Not all of these projections directly regulate feeding[5,6] — some therefore probably have other roles. Alhadeff and colleagues systematically activated AgRP projections to seven brain regions, to search for the projections that mediate the neurons' pain-relieving effect during inflammation. They found a powerful reduction in inflammatory pain following stimulation of AgRP-cell projections to a single target region in the hindbrain, the parabrachial nucleus (PBN). This structure is part of a central pain-processing circuit that relays pain signals from the spinal cord to various brain regions. Notably, the neurons that receive AgRP inputs, which are found in the lateral portion of the PBN (the lPBN), are activated by painful stimuli[7] and inhibited during feeding[7,8]. Presumably, then, lPBN neurons act to suppress appetite in threatening conditions, when eating might be dangerous, whereas their inhibition by input from AgRP neurons supports feeding in conditions of inflammatory pain.

AgRP neurons produce three neurotransmitter molecules that stimulate feeding[9]: AgRP itself, γ-aminobutyric acid (GABA) and neuropeptide Y (NPY). Such co-transmission of signals by multiple molecules is widespread in the brain, but breaking down co-transmission into its constituent parts to understand its functions is challenging. Alhadeff *et al.* overcame this challenge, investigating which of the three molecules were essential for the pain-inhibiting effect of AgRP neurons by injecting each neurotransmitter into the lPBN. Neither AgRP

nor GABA had a pain-relieving effect. But NPY suppressed inflammatory pain by acting through the Y1 receptor on lPBN neurons.

Finally, the authors demonstrated that acute pain led to a sharp decrease in the activity of AgRP neurons. A similar decrease in AgRP activity occurs when an animal first senses food[10], and this change in activity is thought to be important for the termination of further food seeking and a transition to food intake, which is then positively reinforced by structures in the hypothalamus other than the arcuate nucleus[11]. Taking this together with the authors' data, a picture emerges in which acute pain prompts a behavioural transition by suppressing the activity of AgRP neurons. This inhibition prevents the AgRP cells from activating downstream brain regions involved in feeding, and enables pain signals from the spine to spread from the lPBN to other brain regions, indicating the need to avoid noxious stimuli (Fig. 1a).

By contrast, inflammatory pain does not require rapid behavioural responses and is filtered out by active AgRP cells, which might reduce the activity of lPBN neurons to prevent spreading of pain information to other brain regions and so maintain food seeking (Fig. 1b). This previously unknown mechanism for the management of competing needs provides insights into how hypothalamic computations use both the neurochemical properties and the connectivity of neural circuits to make adaptive decisions about behaviour.

Alhadeff and colleagues' work has several implications. First, it provides evidence that the potency of AgRP-mediated long-term pain relief is comparable to that of opiates — at least, in the authors' long-term pain test. As they point out, differences in the processing of

acute and chronic pain suggest that treatments for the two should be aimed at different target cells or proteins. In addition, designing painkillers that lack the off-target effects of opiates is desirable. Alhadeff and colleagues point to NPY–Y1-receptor signalling in the lPBN as a potential site of action for chronic painkillers.

Second, the authors' thorough characterization of a pathway in which signals for two negative states (hunger and pain) interact paves the way to understanding the biological mechanisms that define other complex and dynamic hierarchies in human and animal behaviours. Similar principles at work in other brain regions might further support unaltered food intake during inflammatory pain in hungry mice — for example, by promoting pain inhibition during meals. This painkilling effect probably would not rely on AgRP cells, because their activity is reduced after the sensory detection of food[10], so the behavioural hierarchy at work during feeding itself is probably controlled by other neuronal populations. One possibility is that this hierachy is mediated by neurons in the lateral hypothalamus, which is connected to the PBN (ref. 12) and contains several groups of neurons that are active during feeding[13–15].

As another example, concurrent negative states of hunger and food aversion contribute to eating disorders such as anorexia nervosa: food-related cues elicit aversion, impairing food intake. Delineating interactions between the neurons that mediate hunger and those that control emotional responses to food could shed light on the mechanisms underlying eating disorders. ■

Alexey Ponomarenko *is at the Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich-Heine-University, Düsseldorf 40225, Germany.* **Tatiana Korotkova** *is at the Max Planck Institute for Metabolism Research, Cologne 50931, Germany.*
e-mails: alexey.ponomarenko@med.uni-duesseldorf.de; tatiana.korotkova@sf.mpg.de

1. Alhadeff, A. L. *et al. Cell* **173**, 140–152 (2018).
2. Aponte, Y., Atasoy, D. & Sternson, S. M. *Nature Neurosci.* **14**, 351–355 (2011).
3. Gropp, E. *et al. Nature Neurosci.* **8**, 1289–1291 (2005).
4. Luquet, S., Perez, F. A., Hnasko, T. S. & Palmiter, R. D. *Science* **310**, 683–685 (2005).
5. Betley, J. N., Cao, Z. F. H., Ritola, K. D. & Sternson, S. M. *Cell* **155**, 1337–1350 (2013).
6. Steculorum, S. M. *et al. Cell* **165**, 125–138 (2016).
7. Campos, C. A., Bowen, A. J., Roman, C. W. & Palmiter, R. D. *Nature* **555**, 617-622 (2018).
8. Carter, M. E., Soden, M. E., Zweifel, L. S. & Palmiter, R. D. *Nature* **503**, 111–114 (2013).
9. Krashes, M. J., Shah, B. P., Koda, S. & Lowell, B. B. *Cell Metab.* **18**, 588–595 (2013).
10. Chen, Y., Lin, Y.-C., Kuo, T.-W. & Knight, Z. A. *Cell* **160**, 829–841 (2015).
11. Sternson, S. M. & Eiselt, A.-K. *Annu. Rev. Physiol.* **79**, 401–423 (2017).
12. Saper, C. B. & Loewy, A. D. *Brain Res.* **197**, 291–317 (1980).
13. Jennings, J. H. *et al. Cell* **160**, 516–527 (2015).
14. Nieh, E. H. *et al. Cell* **160**, 528–541 (2015).
15. Carus-Cadavieco, M. *et al. Nature* **542**, 232–236 (2017).

ARTICLE

Quaternary stereocentres via an enantioconvergent catalytic S$_N$1 reaction

Alison E. Wendlandt[1], Prithvi Vangal[1] & Eric N. Jacobsen[1]*

The unimolecular nucleophilic substitution (S$_N$1) mechanism features prominently in every introductory organic chemistry course. In principle, stepwise displacement of a leaving group by a nucleophile via a carbocationic intermediate enables the construction of highly congested carbon centres. However, the intrinsic instability and high reactivity of the carbocationic intermediates make it very difficult to control product distributions and stereoselectivity in reactions that proceed via S$_N$1 pathways. Here we report asymmetric catalysis of an S$_N$1-type reaction mechanism that results in the enantioselective construction of quaternary stereocentres from racemic precursors. The transformation relies on the synergistic action of a chiral hydrogen-bond-donor catalyst with a strong Lewis-acid promoter to mediate the formation of tertiary carbocationic intermediates at low temperature and to achieve high levels of control over reaction enantioselectivity and product distribution. This work provides a foundation for the enantioconvergent synthesis of other fully substituted carbon stereocentres.

Quaternary stereogenic centres are important structural motifs in natural products and biologically active compounds, conferring valuable structural, conformational and metabolic properties. Their construction has long been recognized as an important challenge to the field of chemical synthesis, and several distinct catalytic, enantioselective approaches have been developed in response[1–4]. Notable examples include cycloadditions[5], α- and β-alkylation and arylation of carbonyls[6–8], 3,3'-additions[9], S$_N$2' reactions[10] and Heck-type
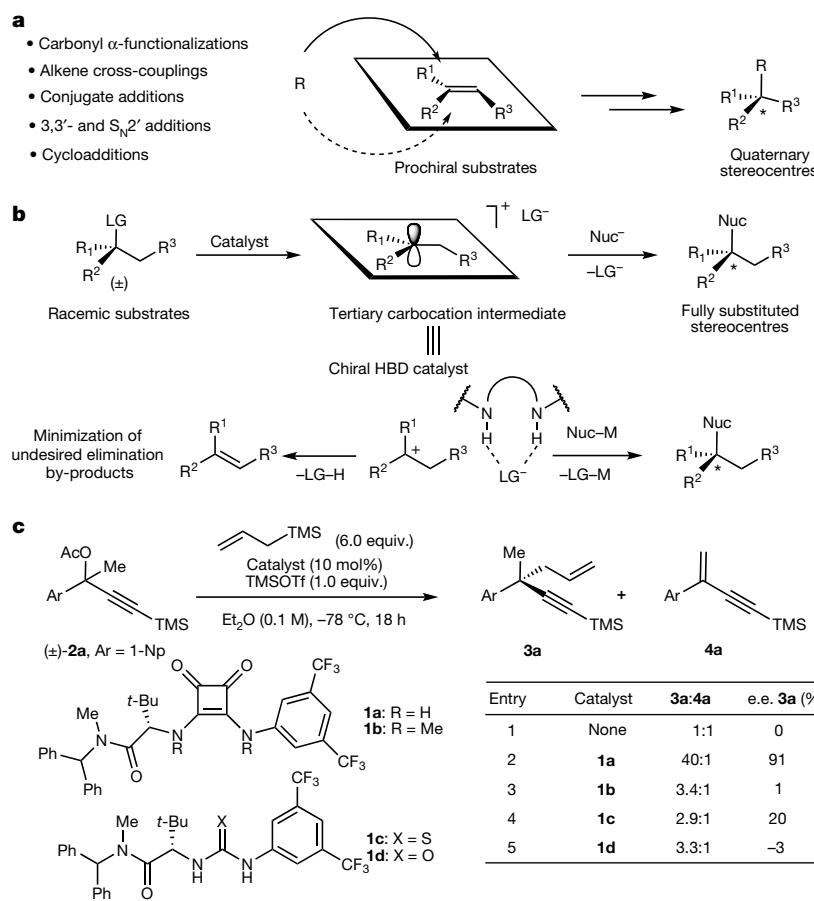

Fig. 1 | Approaches to the enantiocontrolled construction of quaternary stereocentres.
a, Traditional methods for the synthesis of quaternary-stereocentre-containing molecules use stereochemically defined prochiral substrates.
b, The S$_N$1 approach to the construction of quaternary stereocentres described here. **c**, Enantioselective allylation of propargyl acetates using chiral squaramide catalysts and TMSOTf as a promoter. LG, leaving group; Np, naphthyl; TMS, trimethylsilyl.

Entry	Catalyst	3a:4a	e.e. 3a (%)
1	None	1:1	0
2	1a	40:1	91
3	1b	3.4:1	1
4	1c	2.9:1	20
5	1d	3.3:1	–3

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. *e-mail: jacobsen@chemistry.harvard.edu
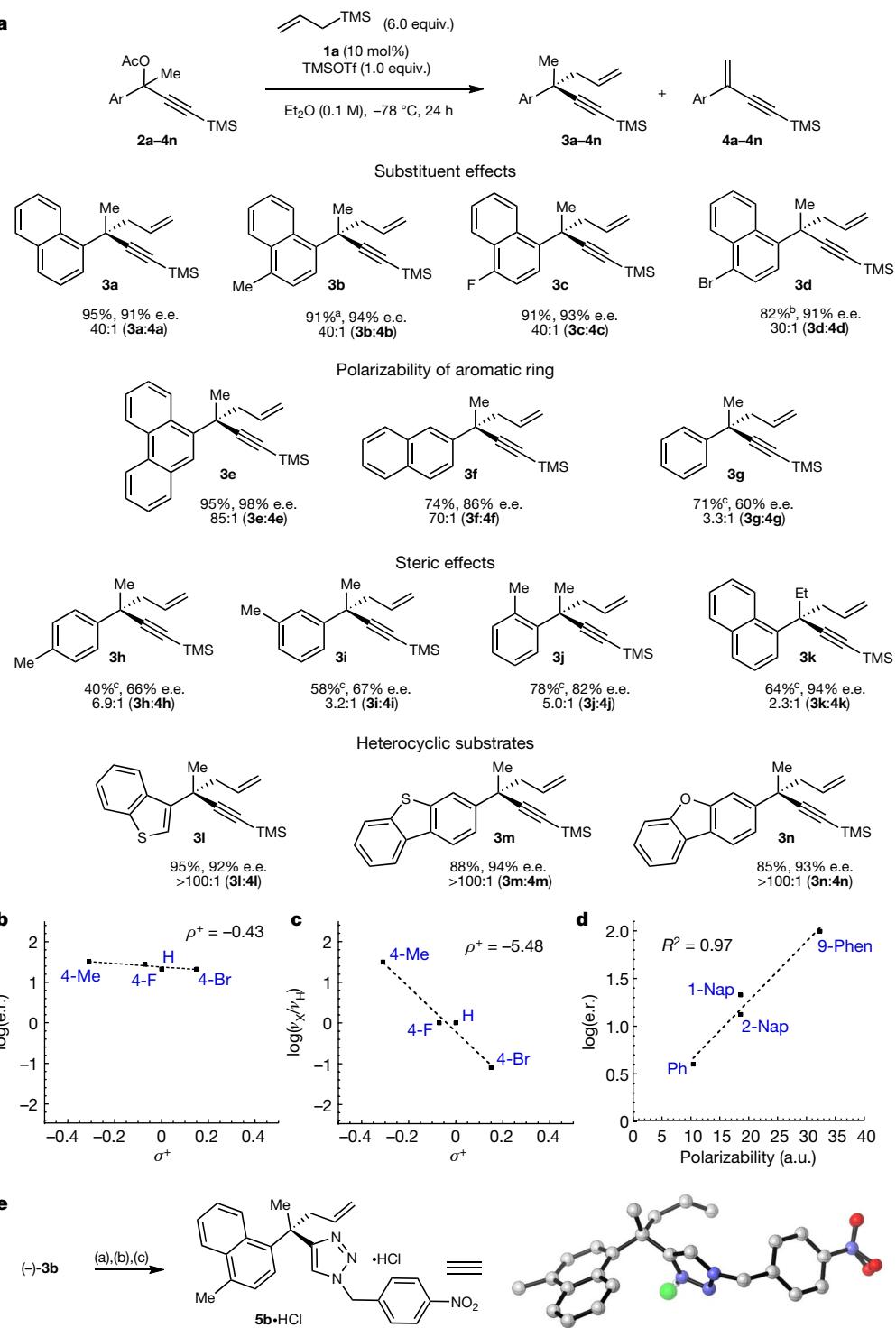
Fig. 2 | Asymmetric allylation of propargyl acetates. a, Substrate scope. Reactions were run on a 0.6-mmol scale with 0.1 equiv. **1a**, 1.0 equiv. TMSOTf and 6.0 equiv. allyltrimethylsilane in 0.1 M Et$_2$O at −78 °C for 24 h. [a]Reaction time was 4 h. [b]Reaction time was 14 d. [c]NMR yield. **b, c**, Hammett plot of the σ^+ values of the substituents in **2a–2d** versus the enantiomeric ratios (log(e.r.)) obtained in the formation of **3a–3d** (**b**) ($\rho^+ = -0.43$) and versus the relative reaction rates (log(ν_X/ν_H)) determined for each substrate (**c**) ($\rho^+ = -5.48$). **d**, Linear free-energy plot of the calculated polarizability of the aromatic rings in **2a** and **2e–2g** versus the enantiomeric ratios (log(e.r.)) obtained in the formation of **3a** and **3e–3g** ($R^2 = 0.97$). a.u., arbitrary units. **e**, The absolute configuration of (−)-**3b** was determined by X-ray crystallography (structure shown), following derivatization to triazole **5b**; the configuration of all other products was assigned by analogy. Conditions: (a) tetra-*n*-butylammonium fluoride (2.0 equiv.), THF, room temperature; (b) 4-nitrobenzylbromide (1.1 equiv.), NaN$_3$ (1.1 equiv.), CuSO$_4$ (0.1 equiv.), sodium ascorbate (0.2 equiv.), *t*BuOH/H$_2$O (1:2), 50 °C; (c) HCl (3 M in Et$_2$O).

cross-couplings[11]. Each of these methods relies on enantiofacial addition across a prochiral substrate (Fig. 1a) and therefore requires the preparation of stereochemically well-defined starting materials (such as trisubstituted olefins) and subsequent enantioselective bond formation.

We envisioned that stepwise nucleophilic substitution reactions that proceed through prochiral carbocationic intermediates could provide a useful and complementary strategy for the enantioselective synthesis of compounds with quaternary stereocentres. Unlike the synthetic approaches noted above, quaternary-stereocentre construction via an S$_N$1-like pathway might be stereoablative[12] and could therefore use readily accessed racemic compounds as substrates (Fig. 1b). Although realization of this strategy would lift the requirement for stereocontrol in the synthesis of the substrate, doing so requires several

very substantial challenges to be overcome. The requisite catalytic system must (a) generate a reactive tertiary carbocationic intermediate, (b) minimize undesired elimination and rearrangement pathways, and (c) exert enantiocontrol in additions of a carbon-centred nucleophile to a high-energy cationic intermediate. As a result, despite the practical appeal of an enantioconvergent approach to the construction of quaternary stereocentres, only isolated examples have been reported so far[9,13,14].

Over the past decade, chiral, dual hydrogen-bond-donor (HBD) catalysts have been developed that promote enantioselective nucleophilic substitution and addition reactions via ion-pair intermediates. These catalysts promote ion-pair formation via direct anion abstraction[15] or by substrate protonation with a co-catalytic Brønsted acid[16].

a

b
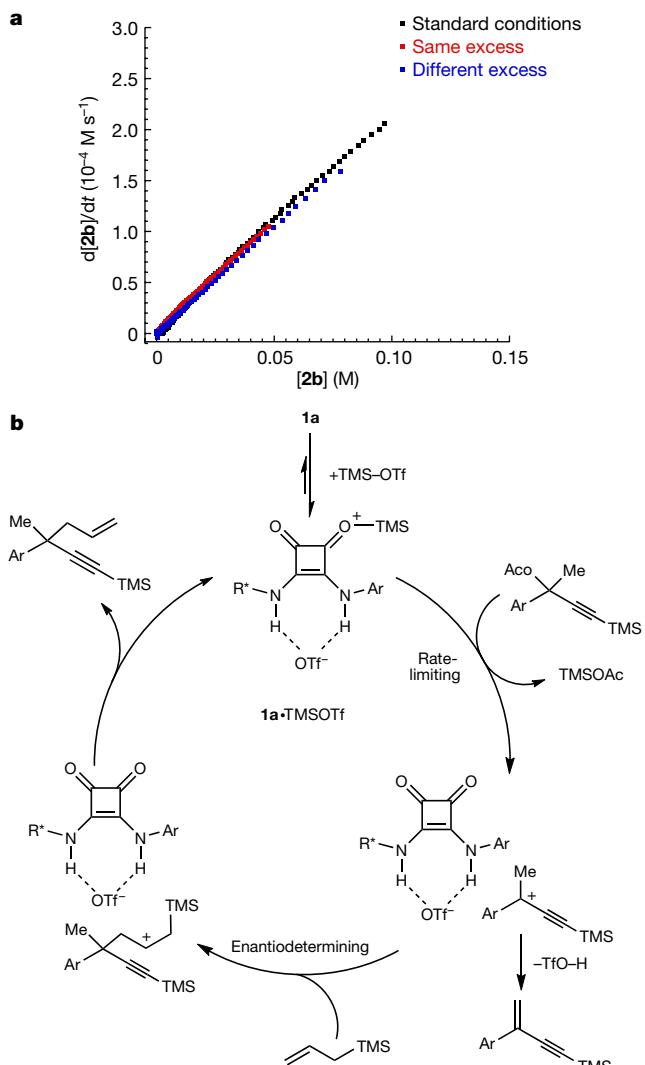
Fig. 3 | Kinetic data and catalytic cycle. a, Reaction-progress kinetic analysis of the reaction of **2b** with allyltrimethylsilane ([**2b**], concentration of **2b**). Standard conditions: [allyltrimethylsilane]$_0$ = 0.25 M, [**2b**]$_0$ = 0.10 M; 'same excess' conditions: [allyltrimethylsilane]$_0$ = 0.195 M, [**2b**]$_0$ = 0.047 M; 'different excess' conditions: [allyltrimethylsilane]$_0$ = 0.315 M, [**2b**]$_0$ = 0.08 M; [X]$_0$, initial concentration of X. **b**, Proposed catalytic mechanism for the enantioselective allylation of propargyl acetates.

Asymmetric induction is typically achieved from the resultant ion pair as a consequence of specific attractive non-covalent interactions between the corresponding cationic intermediate and the chiral HBD catalyst[17–19]. Reported examples have been limited to heteroatom-stabilized cations, owing to the challenges in generating the requisite ion pair and suppressing elimination and rearrangement pathways. The ability of HBD catalysts to control enantioselective nucleophile addition into non-heteroatom-stabilized carbocations has, to our knowledge, not been demonstrated.

It was discovered recently that chiral squaramide catalysts could be used in conjunction with Lewis acids such as trimethylsilyl trifluoromethanesulfonate (TMSOTf) to promote enantioselective reactions[20]. This dual-catalyst system was shown to promote the formation of oxocarbenium ions from dialkyl acetals—substrates that are unreactive under previously developed HBD-promoted reaction manifolds—while still engaging in attractive non-covalent interactions to achieve enantioinduction. We envisioned that the strong ionizing ability of this dual-catalyst system could provide access to carbocationic intermediates that lack heteroatom stabilization, thus allowing us to examine whether small-molecule HBDs can be used to promote

productive, enantioselective reaction pathways from such high-energy intermediates.

Reaction development

After an extensive evaluation of potential tertiary electrophile–carbon-centred nucleophile coupling partners, the reaction of propargyl acetate **2a** with allyltrimethylsilane was identified as a useful model system with which to test this proposal (Fig. 1c). In the absence of an HBD catalyst, the Lewis-acid-promoted reaction affords a 1:1 mixture of the desired product **3a** and the elimination product **4a** (Fig. 1c, entry 1). When readily accessed squaramide **1a** (10 mol%) was added to the reaction, however, **3a** was obtained in high yield (40:1 **3a**:**4a**; Fig. 1c, entry 2) and enantioselectivity (91% enantiomeric excess, e.e.). Product ratio and enantioselectivity were strongly dependent on the nature of the HBD moiety: the related N,N-dimethylated squaramide (**1b**), thiourea (**1c**) and urea (**1d**) catalysts afforded **3a** in low yield and enantiomeric excess (Fig. 1c, entries 3–5). No reaction was observed with squaramide, thiourea or urea HBD catalysts in the absence of TMSOTf.

We then evaluated a series of tertiary propargyl acetates to probe the reaction scope and to generate preliminary information about the mechanism of the enantioselective substitution reaction (Fig. 2a). Substrates with electron-donating (such as **2b** and **2c**) and electron-withdrawing (**2d**) substituents underwent allylation with high enantioselectivity (>90% e.e.) and product selectivity (>30:1 **3**:**4**). A linear correlation with a small negative slope ($\rho^+ = -0.43$, Fig. 2b) was observed between the Hammett substituent σ^+ constants and log(e.r.) for substrates **2a**–**2d** (e.r., enantiomeric ratio). By contrast, a linear correlation with a large negative slope ($\rho^+ = -5.48$, Fig. 2c) was obtained from the corresponding plot of the σ^+ constants versus log(ν_X/ν_H) for the same substrates (where ν_X/ν_H is the reaction rate of substrate that contains substituent X relative to that of the analogous unsubstituted substrate). The observation of a linear free-energy dependence (ρ^+) of this magnitude provides direct evidence of positive charge accumulation in the rate-determining transition state, consistent with an S$_N$1-type ionization mechanism[21].

Despite the very subtle dependence of enantiomeric excess on the electronic properties of the substrate substituents noted above, reaction enantioselectivity was strongly responsive to changes in the expanse and position of the aryl moiety of the substrate. A linear correlation was observed between polarizability values calculated for the aryl substituent[22] and log(e.r.) of products **3a** and **3e**–**3g** (Fig. 2d), indicating that stabilizing aromatic interactions are likely to serve as a contributing factor in enantiodifferentiation[23]. Evidence for the existence of such stabilizing interactions could be gleaned from computational analysis of the putative complex between catalyst **1a** and substrate **2a** (Supplementary Fig. 9). Steric congestion near the reaction site also correlates with enantioselectivity. Thus, the o-tolyl-substituted derivative **2j** underwent an allylation reaction to afford a product with higher enantiomeric excess (82%) than obtained using the p- or m-substituted analogues **2h** and **2i** (66%–67%). Similarly, the ethyl-substituted product **3k** was obtained in higher enantiomeric excess (94%) than the methyl-substituted product **3b** (91%).

Substrates containing electron-rich heterocycles also underwent highly enantioselective substitution. Representative S- and O-heterocyclic substrates underwent reaction with allyltrimethylsilane to afford quaternary products (**3l**–**3n**) in high yield and enantiomeric excess and with no detectable elimination by-products. Following derivatization, the absolute stereochemistry of product **3b** was determined by using X-ray crystallography (Fig. 2e).

Mechanistic studies

We undertook a mechanistic study of the reaction between a representative tertiary propargyl acetate substrate and allyltrimethylsilane promoted by squaramide **1a** and TMSOTf to obtain insights into the underlying catalytic mechanism. The disappearance of **2b** could be monitored over the entire course of the reaction using in situ infrared

Fig. 4 | Mechanistic studies to probe the post-rate-limiting steps of the allylation reaction. a, Crossover experiment to establish the irreversible formation of the alkene by-product (illustrated here for **4a**). **b**, Partial reaction with scalemic **2f**, demonstrating that allylation proceeds via a stereoablative mechanism rather than by a dynamic kinetic-Resolution process. **c**, Predicted and measured $^{12}C/^{13}C$ kinetic isotope effects (KIEs) and equilibrium isotope effects (EIEs). The experimentally determined isotope effects (numbers) are consistent with the predicted mechanistic scenario (top row).

spectroscopy. Runs carried out at different initial concentrations of **2b** but with the same excess in concentration of allyltrimethyl silane (a 'same-excess' experiment[24]) produce good graphical overlay in the kinetic data (Fig. 3a), demonstrating that no catalyst decomposition or product inhibition occurs over the course of the reactions. The linearity of the plot further indicates that the reaction obeys a first-order rate dependence overall. Runs carried out with a different excess in the initial concentration of allyltrimethyl silane relative to **2b** also produce good overlay in the kinetic data (a 'different-excess' experiment[24]), revealing that the reaction obeys a first-order rate dependence on the concentration of **2b** and has no rate dependence on the concentration of allyltrimethylsilane. These kinetic findings are consistent with a stepwise reaction mechanism whereby substrate C–O cleavage is turnover-limiting and nucleophile addition occurs in a post-turnover-limiting step (Fig. 3b). Kinetic studies further revealed a sub-first-order dependence of the reaction rate on the concentration of TMSOTf, and a first-order dependence of the reaction rate on the concentration of **1a** with a non-zero y intercept. The kinetic dependence on the concentrations of **1a** and TMSOTf is consistent with pre-equilibrium formation of a resting-state **1a**·TMSOTf complex that

reacts directly with substrate **2**. The non-zero y intercept is consistent with the competing background reaction that is observed in the absence of **1a** (see Supplementary Information). The observation that optimal enantioselectivities are obtained under conditions where a background, uncatalysed reaction is expected is intriguing, and the subject of continued study.

Having established that the squaramide-TMSOTf-promoted formation of the carbocationic intermediate is rate-limiting, we performed a series of experiments to interrogate the critical post-rate-limiting steps. We determined that the formation of the elimination by-product was irreversible on the basis of a crossover experiment in which 1-naphthyl-substituted enyne (**4a**, 0.25 equiv.) was introduced to the reaction of 2-naphthyl-substituted acetate (**2f**) under otherwise standard reaction conditions. This reaction afforded 2-naphthyl-substituted product **3f** in 80% yield, and alkene **4a** was recovered in 97% yield with no trace of 1-naphthyl allylated product **3a** detected (Fig. 4a).

To evaluate whether the reaction proceeds through an enantio-selective or enantiospecific mechanism, the allylation was carried out by subjecting scalemic substrate (−)-**2f** (81% e.e.) to both enantiomers of the squaramide catalyst **1a** (Fig. 4b). After 1 h of reaction time,

product **3f** was obtained in 86% enantiomeric excess and 24% yield using (S)-**1a**; in the presence of (R)-**1a**, product **3f** was obtained in similar yield but with opposite enantioselectivity (−85% e.e.). In both cases, the substrate **2f** that was recovered was observed to have undergone only a small degree of epimerization, comparable to that observed when **2f** was treated with TMSOTf and in the absence of squaramide catalyst. The results of these experiments are consistent with a stereoablative mechanism, that is, an enantioselective process that proceeds through an achiral carbocationic intermediate. By contrast, a dynamic kinetic-Resolution pathway can be ruled out, whereby **2f** undergoes rapid racemization and one enantiomer preferentially undergoes stereospecific substitution.

We considered two limiting mechanistic possibilities with regard to the enantiodetermining step: (a) irreversible nucleophile addition followed by rapid silyl elimination (Fig. 4c, top), and (b) rapid and reversible nucleophile addition, followed by enantiodetermining silyl elimination (Fig. 4c, bottom). These two scenarios are predicted to produce different carbon isotope effects at the allyl fragment. The carbon kinetic isotope effects (KIEs) were determined with natural-abundance materials using an NMR methodology[25] (Fig. 4c, see Supplementary Information). A large primary KIE of 1.027 was observed at the position of bond formation (internal allylic methylene), whereas no KIE was observed at the terminal position. These results demonstrate that the first C–C bond-forming step is irreversible and therefore enantiodetermining.

Conclusion

We have shown that the cooperative effect of chiral squaramides and TMSOTf generates tertiary carbocations that lack heteroatom stabilization from racemic precursors, controls enantioselectivity in additions of a carbon-centred nucleophile, and attenuates undesired elimination pathways. The strategy outlined here may be generalizable to the construction of many types of highly congested stereogenic centre.

Data availability

The crystallographic data for compound **5b**·HCl can be obtained free of charge from the Cambridge Crystallographic Data Centre (https://www.ccdc.cam.ac.uk) under identifier CCDC 1822228. The raw data for the kinetics experiments are available from the corresponding author on request. All other data that support these findings are available within the paper or Supplementary Information.

1. Quasdorf, K. W. & Overman, L. E. Catalytic enantioselective synthesis of quaternary carbon stereocentres. *Nature* **516**, 181–191 (2014).
2. Liu, Y., Han, S.-J., Liu, W.-B. & Stoltz, B. M. Catalytic enantioselective construction of quaternary stereocenters: assembly of key building blocks for the synthesis of biologically active molecules. *Acc. Chem. Res.* **48**, 740–751 (2015).
3. Das, J. P. & Marek, I. Enantioselective synthesis of all-carbon quaternary stereogenic centers in acyclic systems. *Chem. Commun.* **47**, 4593–4623 (2011).
4. Feng, J., Holmes, M. & Krische, M. J. Acyclic quaternary carbon stereocenters via enantioselective transition metal catalysis. *Chem. Rev.* **117**, 12564–12580 (2017).
5. Wilson, R. M., Jen, W. S. & MacMillan, D. W. C. Enantioselective organocatalytic intramolecular Diels−Alder reactions. The asymmetric synthesis of solanapyrone D. *J. Am. Chem. Soc.* **127**, 11616–11617 (2005).
6. Krautwald, S., Sarlah, D., Schafroth, M. A. & Carreira, E. M. Enantio- and diastereo-divergent dual catalysis: α-allylation of branched aldehydes. *Science* **340**, 1065–1068 (2013).
7. Behenna, D. C. & Stoltz, B. M. The enantioselective Tsuji allylation. *J. Am. Chem. Soc.* **126**, 15044–15045 (2004).
8. Murphy, J. J., Bastida, D., Paria, S., Fagnoni, M. & Melchiorre, P. Asymmetric catalytic formation of quaternary carbons by iminium ion trapping of radicals. *Nature* **532**, 218–222 (2016).
9. Zhang, P., Le, H., Kyne, R. E. & Morken, J. P. Enantioselective construction of all-carbon quaternary centers by branch-selective Pd-catalyzed allyl–allyl cross-coupling. *J. Am. Chem. Soc.* **133**, 9716–9719 (2011).
10. Jung, B. & Hoveyda, A. H. Site- and enantioselective formation of allene-bearing tertiary or quaternary carbon stereogenic centers through NHC–Cu-catalyzed allylic substitution. *J. Am. Chem. Soc.* **134**, 1490–1493 (2012).
11. Mei, T.-S., Patel, H. H. & Sigman, M. S. Enantioselective construction of remote quaternary stereocentres. *Nature* **508**, 340–344 (2014).
12. Bhat, V., Welin, E. R., Guo, X. & Stoltz, B. M. Advances in stereoconvergent catalysis from 2005 to 2015: transition-metal-mediated stereoablative reactions, dynamic kinetic Resolutions, and dynamic kinetic asymmetric transformations. *Chem. Rev.* **117**, 4528–4561 (2017).
13. Braun, M. & Kotter, W. Titanium(IV)-catalyzed dynamic kinetic asymmetric transformation of alcohols, silyl ethers, and acetals under carbon allylation. *Angew. Chem. Int. Ed.* **43**, 514–517 (2004).
14. Zhao, W., Wang, Z., Chu, B. & Sun, J. Enantioselective formation of all-carbon quaternary stereocenters from indoles and tertiary alcohols bearing a directing group. *Angew. Chem. Int. Ed.* **54**, 1910–1913 (2015).
15. Reisman, S. E., Doyle, A. G. & Jacobsen, E. N. Enantioselective thiourea-catalyzed additions to oxocarbenium ions. *J. Am. Chem. Soc.* **130**, 7198–7199 (2008).
16. Xu, H., Zuend, S. J., Woll, M. G., Tao, Y. & Jacobsen, E. N. Asymmetric cooperative catalysis of strong Brønsted acid-promoted reactions using chiral ureas. *Science* **327**, 986–990 (2010).
17. Brak, K. & Jacobsen, E. N. Asymmetric ion-pairing catalysis. *Angew. Chem. Int. Ed.* **52**, 534–561 (2013).
18. Kennedy, C. R., Lin, S. & Jacobsen, E. N. The cation–π interaction in small-molecule catalysis. *Angew. Chem. Int. Ed.* **55**, 12596–12624 (2016).
19. Neel, A. J., Hilton, M. J., Sigman, M. S. & Toste, F. D. Exploiting non-covalent π interactions for catalyst design. *Nature* **543**, 637–646 (2017).
20. Banik, S. M., Levina, A., Hyde, A. M. & Jacobsen, E. N. Lewis acid enhancement by hydrogen-bond donors for asymmetric catalysis. *Science* **358**, 761–764 (2017).
21. Brown, H. C. & Okamoto, Y. Substituent constants for aromatic substitution. *J. Am. Chem. Soc.* **79**, 1913–1917 (1957).
22. McKinney, J. D., Gottschalk, K. E. & Pedersen, L. The polarizability of planar aromatic systems. An application to polychlorinated biphenyls (PCB's), dioxins and polyaromatic hydrocarbons. *J. Mol. Struct. (Theochem)* **105**, 427–438 (1983).
23. Hunter, C. A. & Sanders, J. K. M. The nature of π–π interactions. *J. Am. Chem. Soc.* **112**, 5525–5534 (1990).
24. Blackmond, D. G. Reaction progress kinetic analysis: a powerful methodology for mechanistic studies of complex catalytic reactions. *Angew. Chem. Int. Ed.* **44**, 4302–4320 (2005).
25. Singleton, D. A. & Thomas, A. A. High-precision simultaneous determination of multiple small kinetic isotope effects at natural abundance. *J. Am. Chem. Soc.* **117**, 9357–9358 (1995).

Author contributions A.E.W. and E.N.J. conceived the work, A.E.W. and P.V. conducted the experiments, E.N.J. directed the research, and A.E.W., P.V. and E.N.J. wrote the manuscript.

Competing interests The authors declare no competing interests.

ARTICLE

Renewing Felsenstein's phylogenetic bootstrap in the era of big data

F. Lemoine[1,2], J.-B. Domelevo Entfellner[3,4], E. Wilkinson[5], D. Correia[1], M. Dávila Felipe[1], T. De Oliveira[5,6] & O. Gascuel[1,7]*

Felsenstein's application of the bootstrap method to evolutionary trees is one of the most cited scientific papers of all time. The bootstrap method, which is based on resampling and replications, is used extensively to assess the robustness of phylogenetic inferences. However, increasing numbers of sequences are now available for a wide variety of species, and phylogenies based on hundreds or thousands of taxa are becoming routine. With phylogenies of this size Felsenstein's bootstrap tends to yield very low supports, especially on deep branches. Here we propose a new version of the phylogenetic bootstrap in which the presence of inferred branches in replications is measured using a gradual 'transfer' distance rather than the binary presence or absence index used in Felsenstein's original version. The resulting supports are higher and do not induce falsely supported branches. The application of our method to large mammal, HIV and simulated datasets reveals their phylogenetic signals, whereas Felsenstein's bootstrap fails to do so.

The bootstrap method is a widely used statistical approach to study the robustness, bias and variability of numerical estimates[1,2]. It involves resampling with replacement from the original dataset to obtain replications of the original estimate and then, typically, computing the variance and distribution of this estimate. In 1985, Joseph Felsenstein proposed the use of the bootstrap to assess the robustness, or repeatability, of phylogenetic trees[3]. Given a sequence alignment and a reference tree inferred from it, the procedure is: (i) resample, with replacement, the sites of the alignment to obtain pseudo-alignments of the same length, (ii) infer pseudo-trees using the same inference method and (iii) measure the support of every branch in the reference tree as the proportion of pseudo-trees containing that branch. The usefulness, simplicity and interpretability of this method has made it extremely popular in evolutionary studies, to the point that it is generally required for publication of tree estimates in a wide variety of domains (molecular biology, genomics, systematics, ecology, epidemiology and so on). As a result, Felsenstein's article has been cited more than 35,000 times and is ranked in the top 100 of the most cited scientific papers of all time[4]. However, the use of Felsenstein's bootstrap has been questioned on biological grounds, notably regarding assumptions of site independence and homogeneity[5]. Furthermore, the statistical meaning of Felsenstein's bootstrap proportions (FBPs) has been the subject of intense debate[6], the main questions being whether FBPs can be seen as the confidence levels of some test and whether or not they are biased[7–10]. Several methods[9,11,12] have been proposed to correct FBP to better agree with standard ideas of confidence levels and hypothesis testing. These works have greatly contributed to the understanding of what Felsenstein's bootstrap is and what it is not. However, FBP correction methods are limited to relatively small datasets for mathematical and computational reasons (for example, double bootstrapping), and the original method is still often used; a Google Scholar search reveals about 2,000 citations of Felsenstein's paper in 2017. As has previously been stated[13], "consensus has been reached among practitioners, if not among statisticians and theoreticians" and "many systematists have adopted Hillis and Bull's "70%" value as an indication of support". The alternatives to FBPs are

the Bayesian posterior probabilities of the tree branches[14]—which are difficult to obtain with large datasets for computational reasons—and the approximate branch supports[15,16], which are computed quickly but provide only a local view. The bootstrap is also computationally heavy, but is easily parallelized and fast algorithms have been designed[17,18].

It is commonly acknowledged[13] that Felsenstein's bootstrap is not appropriate for large datasets that contain hundreds or thousands of taxonomic units (taxa), which are now common as a result of high-throughput sequencing technologies. Though such datasets generally contain a lot of phylogenetic information, the bootstrap proportions tend to be low, especially when the tree is inferred from a single gene, or only a few genes, as illustrated in Fig. 1a with a dataset of approximately 9,000 HIV-1 group M (HIV-1M) *DNA polymerase* (*pol*) sequences. The strongest signal in such a phylogeny generally corresponds to the deep branching of the subtypes. This signal is immediately visible here and is consistent with the common belief regarding subtype branching[19] but some of the HIV-1M subtypes (A, B, D and G) are not supported, and neither is their branching (for example, the grouping of C and H). When using a medium-sized dataset of about 550 randomly selected sequences the FBPs are higher, with most subtypes supported at 70% or more. However, their deep branching is still unresolved (Extended Data Fig. 5).

The reason for such degradation is explained by the core methodology of Felsenstein's bootstrap. A replicated branch must match a reference branch exactly to be accounted for in the FBP value. A difference of just one taxon—which is highly likely to be the case in large datasets—is sufficient for the replicated branch to be counted as absent, even though it is nearly identical to the reference branch[13,20]. There are many biological and computational reasons for the existence of 'rogue' taxa with unstable phylogenetic positions: convergence, recombination, sequence and tree errors, and so on. The standard approach[20–24] is to remove these taxa and relaunch the analysis, but this is statistically questionable and computationally expensive. Furthermore, with a large number of taxa and a low number of sites the phylogenetic signal is weak. The inferred branches are then likely to have errors and a large

[1]Unité Bioinformatique Evolutive, C3BI USR 3756, Institut Pasteur & CNRS, Paris, France. [2]Hub Bioinformatique et Biostatistique, C3BI USR 3756, Institut Pasteur & CNRS, Paris, France. [3]Department of Computer Science, University of the Western Cape, Cape Town, South Africa. [4]South African MRC Bioinformatics Unit, South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa. [5]KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa. [6]Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa. [7]Méthodes et Algorithmes pour la Bioinformatique, LIRMM UMR 5506, Université de Montpellier & CNRS, Montpellier, France. *e-mail: olivier.gascuel@pasteur.fr
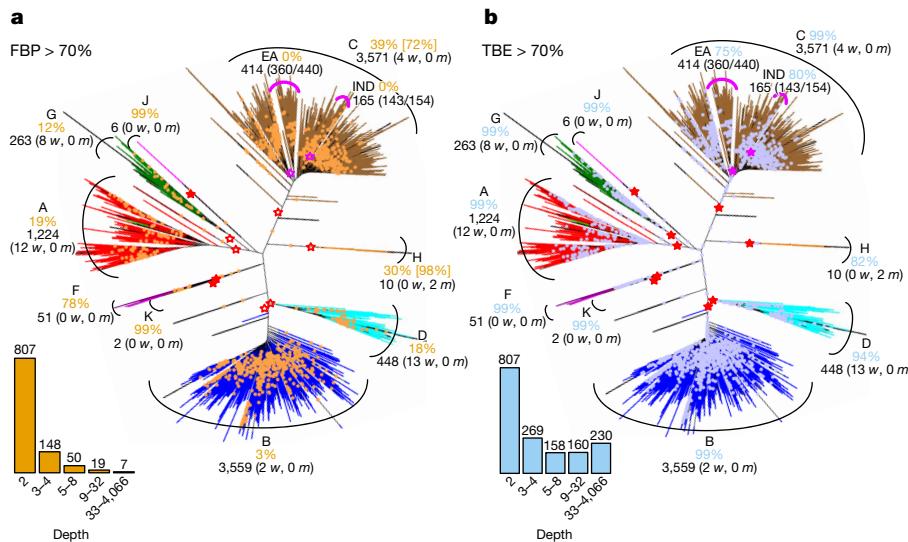
Fig. 1 | FBP and TBE bootstrap supports on the same phylogeny with 9,147 HIV-1M *pol* sequences. a, FBP. **b**, TBE. Subtypes are colourized; recombinant sequences are black; dots correspond to branches with support > 70%[7]. Supports (orange and blue percentages in **a** and **b**, respectively) are given for the tree clades that are closer to the subtypes (red stars, filled when support > 70%); for each of these clades we use jpHMM predictions to provide the number of wrong taxa (*w*) that do not belong to the corresponding subtype, and the number of missing taxa (*m*) that belong to the subtype but not to the clade. The C and the H clades are not supported by FBP but there exist neighbouring clades with FBP

supports larger than 70%, which are shown in square brackets. The same approach is applied to the C sub-epidemics in India (IND) and East Africa (EA). The ratio provides the coverage of the clade: the number of studied taxa (for example, from India) in the clade versus the total number of these taxa in the dataset. The South American clade (SA, included in EA, not shown) is supported by TBE but not by FBP (73% versus 14%, respectively, for 15 taxa in total, with above-defined ratio of 14/14). The histograms provide the number of branches with > 70% support depending on branch depth, which is measured by the number of taxa in the smaller of the two clades defined by the given branch.

fraction of taxa may be unstable, even in the absence of model misspecification of any sort and without long branches.

A statistical approach

Our approach has a simple but sound statistical basis that is partly inspired by Sanderson's monophyly index[25] and partly by our work on gene clusters obtained from expression data[26], both of which are tailored for rooted trees. We replace the branch presence proportion—that is, the expectation of a {0,1} indicator function—of Felsenstein's bootstrap, by the expectation of a refined, gradual function in the [0,1] range, quantifying the branch presence in the bootstrap trees. In doing so, we admit that the inferred branch is not simply correct or incorrect (as with FBP), but that it may contain some errors. Our ultimate aim is to quantify these errors and the presence of the inferred branch in the true tree, using the plug-in principle (see below). We use the transfer distance[27–29], in which the distance $\delta(b,b^*)$ between a branch b of the reference tree T and a branch b^* of a bootstrap tree T^* is equal to the number of taxa that must be transferred (or removed) to make both branches identical (that is, both branches split the set of taxa identically). To measure the presence of b in T^*, we search the branch in T^* that is closest to b and use the 'transfer index', $\phi(b,T^*) = \mathrm{Min}_{b^* \in T^*}\{\delta(b,b^*)\}$.

This index has several important and useful properties. Any branch b splits the taxa into two subsets. If l is the number of taxa and p is the size of the smaller subset induced by b, we have the following properties (see Methods): $\phi(b,T^*) = 0$ if and only if b belongs to T^*; $\phi(b,T^*) \leq p - 1$; $\phi(b,T^*)/(p-1)$ is very close to 1 when T^* is random and l is large (> 100); and $\phi(b,T^*)$ is computed recursively in time proportional to l, just as is FBP. On the basis of these properties, we define the transfer bootstrap expectation (TBE) as:

$$\mathrm{TBE}(b) = 1 - \frac{\overline{\phi(b,T^*)}}{p-1}$$

in which the numerator is the average transfer index among all bootstrap trees. It can easily be seen that TBE ranges from 0 to 1, in which 0 means that the bootstrap trees are random regarding b and 1 means that

b appears in all bootstrap trees. Considering the same set of bootstrap trees, $\mathrm{TBE}(b)$ is necessarily larger than $\mathrm{FBP}(b)$ and the difference is substantial for deep branches, whereas $\mathrm{TBE}(b) = \mathrm{FBP}(b)$ when b defines a (shallow) 'cherry', which is a clade comprising only two taxa (that is, $p = 2$). Importantly, we shall see that TBE supports very few branches showing substantial contradictions with the true tree when used with common thresholds (typically 70%[7] or higher; Fig. 2c, d and Extended Data Figs. 2, 3, 7, 8).

These properties—easy computation, higher supports than FBP and a low number of falsely supported branches—are all highly desirable. Furthermore, TBE has a simple and natural interpretation; for instance, with $l = 1,000$ and $p = 200$, $\mathrm{TBE}(b) = 95\%$ means that, on average, $(200 - 1) \times 0.05 \approx 10$ taxa have to be transferred to recover b in bootstrap replicate trees. This interpretation is radically different from that of FBP, in which b is assessed globally as correct or erroneous. With TBE, branches that are nearly correct are also likely to be supported. Moreover, we can define an instability score for each taxon based on the number of times it is transferred in TBE computations.

TBE uses the same procedure of resampling with replacement as does FBP and thus inherits some of the statistical properties of FBP[3,6,9], as well as the usual properties of the bootstrap method[1,2]. Notably, TBE relies on the same assumptions as FBP regarding site independence and homogeneity, but these assumptions can be relaxed[6], for instance, by using block bootstrapping[30]. Just as with FBP[3], $\mathrm{TBE}(b)$ cannot be interpreted as the probability for the branch b to belong to the true phylogeny. Although deep mathematical approaches[6,9–12,31] have previously been proposed to connect FBP to hypothesis-testing theory, TBE should not be interpreted as the confidence level of some statistical test (with null and alternative hypotheses, distribution of test statistics under the null and so on). TBE is better and more simply interpreted in terms of repeatability: $\mathrm{TBE}(b)$ estimates the extent to which branches identical or similar to b would be recovered when applying the same tree inference method to a new sample of the same size drawn from the same site distribution as the original sample. With large samples, the empirical distribution obtained from observed data comes close to the unknown underlying distribution of this data, and sampling with replacement in the empirical distribution is asymptotically equivalent to drawing
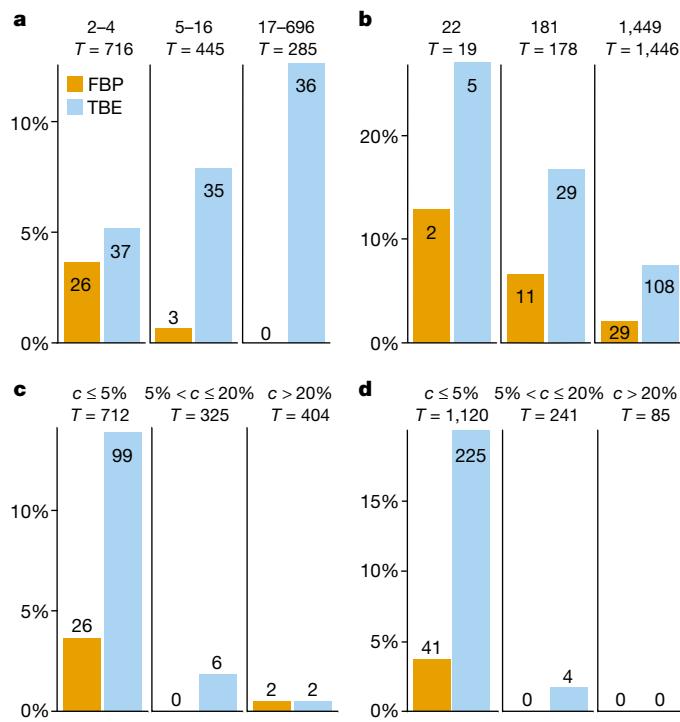
Fig. 2 | FBP and TBE bootstrap supports on the same phylogeny with 1,449 COI-5P mammal sequences using FastTree. The graphs in **a–d** refer to branches with supports > 70%[7], with the vertical axes denoting the percentage of these branches in a given condition (in **b** for example, 22-taxon trees contain 19 internal branches (T), and 2/19 ≈ 10% of branches have FBP > 70%). **a**, Supports regarding branch depth (see definition in Fig. 1). **b**, Supports regarding tree size (that is, number of taxa). **c**, Supports regarding percentage of quartet conflicts (c) with NCBI taxonomy ($c \leq 5\%$, low level of conflict; $5\% < c \leq 20\%$, moderate level of conflict; $c > 20\%$, high level of conflict). **d**, As in **c** but regarding the true tree used for simulations (noisy condition). T, number of internal branches.

samples from the underlying distribution[1,2]. The convergence rate is unknown with models as complex as the ones used in phylogenetics, but our simulation results show that moderate sample sizes suffice to obtain good approximations (Extended Data Fig. 10). When the sample size is extremely large, as in phylogenomic studies using genome-scale sequence alignments[32], both FBP and TBE are expected to be nearly equal to 1 for all branches. Again, this should not be interpreted in terms of absolute truth regarding the phylogenetic inferences, but it simply reflects the closeness of the empirical and underlying distributions and the very small variability of tree estimates. In fact, a high level of repeatability is necessary to trust phylogenetic inferences, but it may be not sufficient. Felsenstein[3] states that the bootstrap "may be misleading if the method used to infer phylogenies is inconsistent". This applies both to FBP and TBE, and is typical for inference methods subject to long-branch attraction. With a consistent, unbiased inference method, we expect the plug-in principle[1,2,6,9] to apply; this principle states that the distribution of the distance between the true tree and the inferred tree can be well-approximated by the distribution of the distance between the inferred and bootstrap trees. Using both real and simulated data, here we show that this principle does apply with maximum-likelihood estimation, a phylogenetic inference method that is typically consistent[33]. In this setting, TBE provides information on the (transfer, quartet-based) distance between the inferred branch and the true tree, and rarely supports poor branches. Moreover, the ability of TBE to identify rogue taxa makes it possible to study them further, to understand why they are phylogenetically unstable and to revise the branch supports.

Analysis of mammal data
We first studied the advantage of using TBE on a large phylogeny of 1,449 mammals, obtained from a usual barcoding marker (COI-5P).
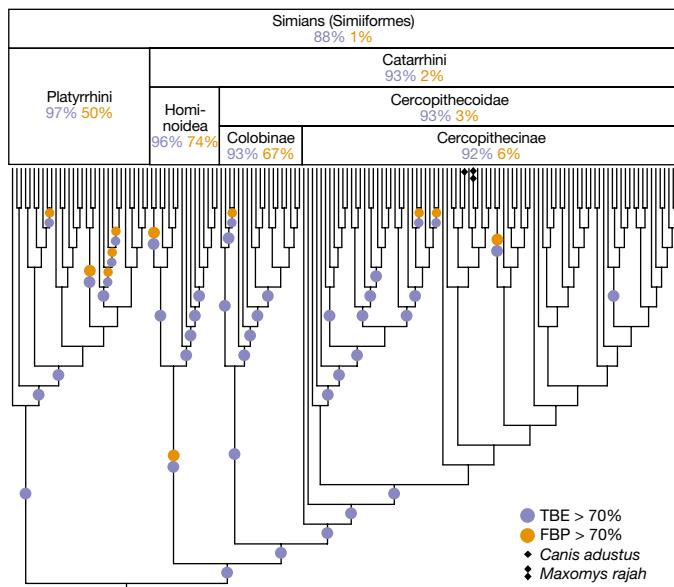

Fig. 3 | FBP and TBE bootstrap supports on the simian clade. The complete tree inferred by FastTree using 1,449 COI-5P mammal sequences is the same as in Fig. 2, but we focus on the simian clade. All simian sequences are included in this clade and two additional non-simian sequences are added, one rogue taxon (*Maxomys rajah*, detected by TBE) and one stable but erroneous taxon with partial sequence (*Canis adustus*); this simian tree is very close to the NCBI taxonomy (< 2.5% of contradicted quartets, when both erroneous taxa are pruned). Platyrrhini, New World monkeys; Cercopithecoidae, Old World monkeys.

The reference and bootstrap trees were inferred by maximum likelihood from the protein alignment (527 sites) using both FastTree[34] and RAxML with rapid bootstrap[17] to check that similar conclusions were drawn with different inference methods. To study the effect of the number of taxa, we randomly selected small- (22 taxa) and medium-sized (181 taxa) datasets and performed the same analyses. The results were compared to the NCBI taxonomy (https://www.ncbi.nlm.nih.gov/taxonomy, accessed April 2016), which represents current thinking about the evolutionary history of mammals. To cope with the low resolution of the NCBI taxonomy, we used a quartet-based topological distance rather than the transfer distance. For all inferred branches, we measured the quartet-based percentage of conflicts with the NCBI taxonomy, and the same approach was used to assess the topological accuracy of FastTree and RAxML phylogenies. As expected in this type of study based on a unique marker, the inferred topologies were relatively poor, and thus challenging for branch support methods. However, RAxML was more accurate than FastTree and had higher branch supports, as is generally observed with rapid bootstrap[16] (Extended Data Figs. 2, 3).

Our results (Fig. 2a–c and Extended Data Figs. 2, 3) indicate clearly that TBE provides some support for deep branches, whereas FBP does not. As expected, the supports for shallow branches are similar between the two methods, and the advantage of TBE is more pronounced with a large number of taxa but still of interest with medium-sized datasets. Comparisons with the NCBI taxonomy show that TBE supports a larger number of weakly contradicted branches than FBP—which fulfils one of the objectives of TBE (nearly correct branches must be supported)—and the number of supported branches with moderate-to-high quartet conflicts remains very low. These results are confirmed by simulations (Fig. 2d and Extended Data Figs. 7, 8).

The advantage of TBE appears clearly when inspecting the tree clades. For example (Fig. 3), the simian clade inferred by FastTree has a strong support with TBE; by contrast, when using FBP, support for this clade is nearly null owing to a large number of rogue taxa in the bootstrap trees, and the same holds true for several sub-clades. The simian clade includes all 152 simian sequences plus two non-simian taxa
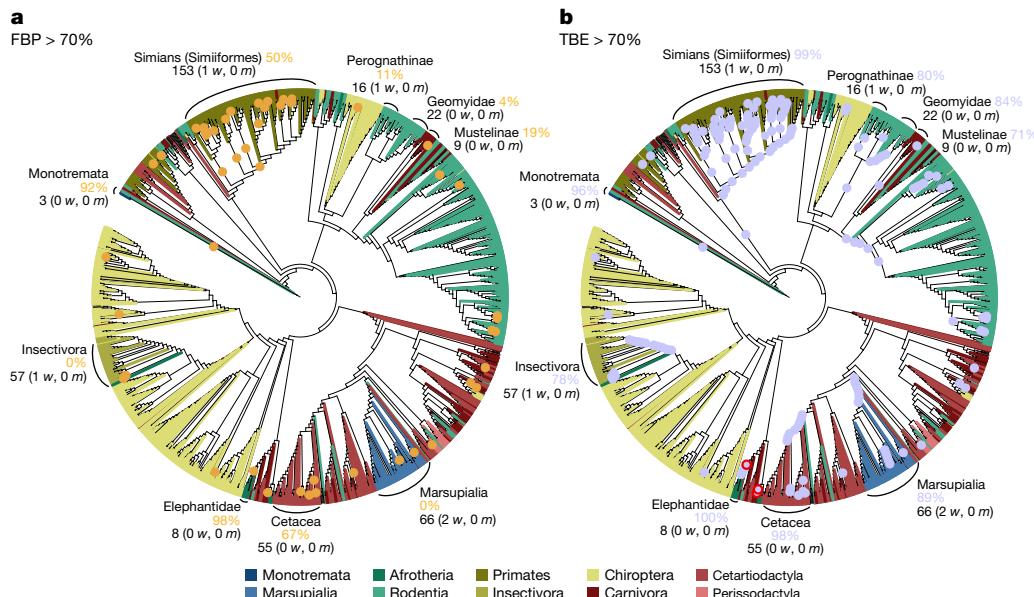
Fig. 4 | FBP and TBE supports on the same phylogeny with 1,449 COI-5P mammal sequences using RAxML with rapid bootstrap. **a**, FBP. **b**, TBE. This phylogeny is more accurate than the one inferred using FastTree (27% versus 38% of contradicted quartets, respectively) but is still relatively poor, especially regarding deep nodes and larger groups. For example, rodents and chiropterans are not monophyletic and are distributed in several subtrees. However, some parts of the tree are more accurate. A few clades are highlighted, corresponding almost exactly to the NCBI taxonomy. For example, all Elephantidae taxa are recovered by RAxML in a single clade that contains only Elephantidae, and Insectivora taxa are included in a clade containing one extra taxon. To select these clades, we minimized the transfer distance with the NCBI taxonomy, in case of ambiguity. See Fig. 1 and Methods for details.

(*Maxomys rajah* and *Canis adustus*). The latter is not a rogue taxon: its sequence is incomplete and very close to the simian sequences for the part that is available, and its position is very stable in the bootstrap trees. By contrast, *Maxomys rajah* is a rogue taxon and is detected as such by TBE (transferred in 659 out of 1,000 bootstrap trees when computing the support of the simian clade). Similar results were found with other well-established clades when using RAxML (Fig. 4). Both FBP and TBE support some small clades, namely the Monotremata and Elephantidae. However, FBP does not support any deep branches, except for the Cetacea (67%) and, to some extent, the simians (50%). TBE provides strong supports for these two groups, but also for five other groups, including the Marsupiala and Insectivora. The latter clade (FBP: 0%, TBE: 78%) contains all Insectivora of the NCBI taxonomy, plus one extra taxon (*Plecotus strelkovi*), which again is detected by TBE as a rogue taxon (transferred in 965 out of 1,000 bootstrap trees). By comparison, the removal of rogue taxa[24] does not substantially improve FBP: eight and three taxa are removed with FastTree and RAxML, respectively, but the number of branches with FBP > 70% remains the same. This is explained by hundreds of taxa, which are relatively unstable but not removed.

Analysis of HIV data
We applied our method to a large dataset of 9,147 HIV-1M *pol* sequences. Datasets of this size are increasingly common in molecular epidemiology and phylodynamics[35]. We retained only sequences that were annotated as non-recombinant by the Los Alamos HIV-1 database using a fast-filtering approach. Among these sequences, 48 recombinant sequences were detected by jpHMM[36]. These 48 sequences were kept in the analyses to study the effect of recombinant sequences, as their presence is inevitable in any HIV dataset. In contrast to that of mammals, the tree topology of HIV-1M strains is essentially unknown. Moreover, it is intrinsically unstable because reconstructing a tree with so many relatively short and possibly recombinant sequences is challenging. Thus, the main expectation is to observe a clear separation between the subtypes. We built the reference and bootstrap trees using FastTree on the DNA sequence alignment (1,043 sites), and performed the same analyses using smaller subsets of 35 and 571 sequences.

Although the deep branching of the subtypes[19] is poorly supported by FBP (Fig. 1a), it becomes apparent with TBE, as when using this approach all subtypes have a support larger than 80% and close to 100% in most cases (Fig. 1b and Extended Data Fig. 5). For example, the subtype B clade (3,559 taxa) has a support of only 3% using FBP, but a support of 99% using TBE. This clade contains all subtype B sequences, plus two taxa detected as recombinant by jpHMM; this means that both supports are likely to be correct insofar as they state that this clade is incorrect (FBP) or nearly correct (TBE). However, FBP fails to detect any phylogenetic signal, whereas TBE reveals that this signal is very strong. The same holds true with other well-described clades. For example, TBE supports the identification of regional variants of HIV-1 subtypes that are of epidemiological importance (such as the East African, Indian and South American subtype C variants), which FBP fails to support. TBE provides a substantial support to a much larger number of deep branches. Again, the advantage of using TBE is higher with large datasets (Extended Data Fig. 4), but is still apparent when using 571-taxon datasets, for which the deep subtype branching and C sub-epidemics are supported by TBE but not FBP (Extended Data Fig. 5). An important feature of TBE is that the supports may be non-local, but attached to 'caterpillar-like' paths, in which the main phylogenetic backbone is connected to a few isolated taxa (Fig. 1b; for example, subtype C). With HIV-1M data, this corresponds to the fact that the subtype roots are usually not well defined owing to recombinant and ancient sequences, which tend to be isolated in basal position. Moreover, the instability score among recombinant sequences is clearly higher than in the sequences that were not detected as recombinant (Extended Data Fig. 6), which supports the biological soundness of the approach and its power to detect recombinant and rogue taxa.

Analysis of simulated data
To check that TBE does not support erroneous branches, we performed extensive computer simulations with various tree sizes and phylogenetic signal levels. We also added unstable taxa that had a weaker phylogenetic signal than the others. The results are highly similar to those obtained using real data, regarding the support of deep branches and the tree size (Extended Data Figs. 7, 8). In all the conditions we

examined, TBE supported very few branches that showed substantial contradictions with the true tree, and the rogue taxa exhibited lower stability (Extended Data Fig. 9). In the absence of rogue taxa (Extended Data Fig. 7), the gain of TBE was still substantial compared to FBP, with almost twice as many branches with support > 70%, thus demonstrating the importance of accounting for the global instability of the inferred tree. Furthermore, we checked the interpretation of TBE as a measure of repeatability (Extended Data Fig. 10) by comparing TBE to its counterpart computed from simulated alignments, rather than bootstrap pseudo-alignments; both simulation- and bootstrap-based supports are highly correlated (Pearson's $\rho = 0.85$) with alignments of moderate length (about 500) and have analogous performance in detecting rogue taxa. Lastly, we checked the validity of the plug-in principle by comparing TBE to the similarity—measured using the normalized transfer index—between the inferred branch and the true tree (Extended Data Fig. 10). Again a high correlation (Pearson's $\rho = 0.74$) was found. When performing the same experiments with FBP similar or slightly lower correlations were observed, probably owing to the discontinuous nature of FBP.

Discussion

The transfer bootstrap thus provides a measure of branch repeatability, or robustness. Our results clearly demonstrate its usefulness, especially with deep branches and large datasets, for which branches known to be essentially correct are supported by TBE but not by FBP. Furthermore, when combined with consistent maximum-likelihood tree estimation, TBE rarely supports poor branches. Importantly, TBE supports are easily interpreted as fractions of unstable taxa. Although our results suggest that 70% is a reasonable threshold from which to start (Extended Data Figs. 2, 3, 4, 8), we suggest that it is better to interpret TBE values depending on the data and the phylogenetic question being addressed; for example, using a lower TBE support threshold with HIV and possibly recombinant sequences, than with mammals. Moreover, our experiments demonstrate the ability of the transfer index to detect unstable taxa responsible for low supports. Lastly, the approach is applicable to rapid bootstrap[17,18] (Fig. 4 and Extended Data Fig. 3) and could be extended to parametric bootstrap[2] and Bayesian branch supports[14].

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0043-0.

1. Efron, B. Bootstrap methods: another look at the jackknife. *Ann. Stat.* **7**, 1–26 (1979).
2. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (Chapman & Hall, New York, 1993).
3. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791 (1985).
4. Van Noorden, R., Maher, B. & Nuzzo, R. The top 100 papers. *Nature* **514**, 550–553 (2014).
5. Sanderson, M. J. Objections to bootstrapping phylogenies: a critique. *Syst. Biol.* **44**, 299–320 (1995).
6. Holmes, S. Bootstrapping phylogenetic trees: theory and methods. *Stat. Sci.* **18**, 241–255 (2003).
7. Hillis, D. M. & Bull, J. J. An empirical test of bootstrapping as a method for assessing confidence in phyogenetic analysis. *Syst. Biol.* **42**, 182–192 (1993).
8. Felsenstein, J. & Kishino, H. Is there something wrong with the bootstrap on phylogenies? A reply to Hillis and Bull. *Syst. Biol.* **42**, 193–200 (1993).
9. Efron, B., Halloran, E. & Holmes, S. Bootstrap confidence levels for phylogenetic trees. *Proc. Natl Acad. Sci. USA* **93**, 7085–7090 (1996).
10. Susko, E. Bootstrap support is not first-order correct. *Syst. Biol.* **58**, 211–223 (2009).
11. Zharkikh, A. & Li, W.-H. Estimation of confidence in phylogeny: the complete- and-partial bootstrap technique. *Mol. Phylogenet. Evol.* **4**, 44–63 (1995).
12. Susko, E. First-order correct bootstrap support adjustments for splits that allow hypothesis testing when using maximum likelihood estimation. *Mol. Biol. Evol.* **27**, 1621–1629 (2010).
13. Soltis, D. E. & Soltis, P. S. Applying the bootstrap in phylogeny reconstruction. *Stat. Sci.* **18**, 256–267 (2003).
14. Huelsenbeck, J. & Rannala, B. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. *Syst. Biol.* **53**, 904–913 (2004).
15. Anisimova, M. & Gascuel, O. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* **55**, 539–552 (2006).
16. Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C. & Gascuel, O. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* **60**, 685–699 (2011).
17. Stamatakis, A., Hoover, P. & Rougemont, J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* **57**, 758–771 (2008).
18. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195 (2013).
19. Hemelaar, J. The origin and diversity of the HIV-1 pandemic. *Trends Mol. Med.* **18**, 182–192 (2012).
20. Sanderson, M. J. & Shaffer, H. B. Troubleshooting molecular phylogenetic analyses. *Annu. Rev. Ecol. Syst.* **33**, 49–72 (2002).
21. Wilkinson, M. Majority-rule reduced consensus trees and their use in bootstrapping. *Mol. Biol. Evol.* **13**, 437–444 (1996).
22. Thorley, J. L. & Wilkinson, M. Testing the phylogenetic stability of early tetrapods. *J. Theor. Biol.* **200**, 343–344 (1999).
23. Thomson, R. C. & Shaffer, H. B. Sparse supermatrices for phylogenetic inference: taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. *Syst. Biol.* **59**, 42–58 (2010).
24. Aberer, A. J., Krompass, D. & Stamatakis, A. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Syst. Biol.* **62**, 162–166 (2013).
25. Sanderson, M. J. Confidence limits on phylogenies: the bootstrap revisited. *Cladistics* **5**, 113–129 (1989).
26. Bréhélin, L., Gascuel, O. & Martin, O. Using repeated measurements to validate hierarchical gene clusters. *Bioinformatics* **24**, 682–688 (2008).
27. Charon, I., Denoeud, L., Guénoche, A. & Hudry, O. Maximum transfer distance between partitions. *J. Classif.* **23**, 103–121 (2006).
28. Day, W. H. E. The complexity of computing metric distances between partitions. *Math. Soc. Sci.* **1**, 269–287 (1981).
29. Lin, Y., Rajan, V. & Moret, B. M. E. A metric for phylogenetic trees based on matching. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **9**, 1014–1022 (2012).
30. Künsch, H. R. The jackknife and the bootstrap for general stationary observations. *Ann. Stat.* **17**, 1217–1241 (1989).
31. Billera, L. J., Holmes, S. P. & Vogtmann, K. Geometry of the space of phylogenetic trees. *Adv. Appl. Math.* **27**, 733–767 (2001).
32. Kumar, S., Filipski, A. J., Battistuzzi, F. U., Kosakovsky Pond, S. L. & Tamura, K. Statistics and truth in phylogenomics. *Mol. Biol. Evol.* **29**, 457–472 (2012).
33. Truszkowski, J. & Goldman, N. Maximum likelihood phylogenetic inference is consistent on multiple sequence alignments, with or without gaps. *Syst. Biol.* **65**, 328–333 (2016).
34. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
35. Grenfell, B. T. et al. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332 (2004).
36. Schultz, A.-K. et al. jpHMM: improving the reliability of recombination prediction in HIV-1. *Nucleic Acids Res.* **37**, W647–W651 (2009).

Author contributions O.G. designed the research; F.L., J.-B.D.E., M.D.F. and O.G. performed the research; F.L. and J.-B.D.E. implemented the algorithms; F.L. and D.C. realized the website and GitHub repositories; F.L. performed the analyses and graphics, with the help of E.W. and T.D.O. for HIV; O.G. wrote the paper with the help of all co-authors.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0043-0.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0043-0.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to O.G.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Definitions and properties of transfer distance and index. The transfer distance[27], also called R-distance[28], was introduced to compare partitions in cluster analysis. In this context, the transfer distance is equal to the minimum number of elements to be transferred (or removed) to transform one partition into the other. Tree branches are commonly seen as bipartitions or splits, as a branch divides the taxa into two subsets situated on its two sides. The most used topological distance between two trees is the Robinson–Foulds distance[37], which is equal to the number of bipartitions that belong to one tree but not the other. The bipartition distance is overly sensitive to some small tree changes, possibly involving a unique taxon[29]. Previous authors[29] have proposed using the transfer distance and designed algorithms to compute a more robust 'matching' distance between trees; although a different task, this is related to the aim of this article. In the following, we first provide basic definitions—following the standard text book for phylogenetic trees[38]—and then demonstrate the properties of the transfer distance in a bootstrap context.

Let X be a fixed set of l taxa. An X-tree is a phylogenetic tree with l leaves labelled by the taxa of X. All reference and bootstrap trees discussed here are X-trees, meaning that they are labelled by the same set of l taxa. Any branch of an X-tree defines a bipartition of X, and the topology of an X-tree can be recovered from its bipartition set. Thus, we will use the terms branch and bipartition to mean the same thing in different contexts. Any bipartition, b, of X can be encoded by a $\{0,1\}$ vector $\mathbf{v}(b)$ of length l, in which the taxa on the same side of the bipartition are encoded by the same value. Note that b is also encoded by $\bar{\mathbf{v}}(b)$, the negation of $\mathbf{v}(b)$ (that is, the zeros are turned into ones, and vice versa). Moreover, the smaller of the two subsets induced by a bipartition b will be called here the 'light side' of b, and p will denote the size of the light side of b ($p \leq l - p$). A bipartition is 'trivial' when it has a unique taxon in its light side ($p = 1$). An X-tree defines l trivial bipartitions corresponding to each of the taxa. These trivial bipartitions are contained in every X-tree, while the other non-trivial bipartitions define the core of the tree topology and are the central subject of phylogenetic studies.

The transfer distance $\delta(b,b^*)$ between a bipartition b of the reference tree T and a bipartition b^* of a bootstrap tree T^* is equal to the number of taxa that must be transferred (or removed) to make both bipartitions identical. The transfer distance is easily defined and computed using the Hamming distance, H, between $\mathbf{v}(b)$ and $\mathbf{v}(b^*)$:

$$\delta(b,b^*) = \text{Min}\{H(\mathbf{v}(b),\mathbf{v}(b^*)), H(\bar{\mathbf{v}}(b),\mathbf{v}(b^*))\}$$

To measure the presence of b in T^*, we search the bipartition in T^* that is closest to b and use the transfer index $\phi(b,T^*) = \text{Min}_{b^* \in T^*}\{\delta(b,b^*)\}$. Based on above definitions, $\delta(b,b^*) = 0$ if and only if $\mathbf{v}(b)$ and $\mathbf{v}(b^*)$ define the same bipartition of X. Thus, the transfer index $\phi(b,T^*) = 0$ if and only if $b \in T^*$. Moreover, let b be any given bipartition of T and t be a taxon on the light side of b. The trivial bipartition $b^* = \{t\}|X - \{t\}$ is found in any bootstrap tree T^* and $\delta(b,b^*) = p - 1$. There may well be another bipartition closer to b in T^*, but at least this ensures that $\phi(b,T^*) \leq p - 1$, and thus the transfer support, TS, satisfies:

$$\text{TS}(b,T^*) = 1 - \frac{\phi(b,T^*)}{p-1} \in [0,1]$$

and $\text{TS}(b,T^*) = 1$ if and only if $b \in T^*$. Let $\mathbf{1}_b(T^*)$ be the indicator function equal to 1 when $b \in T^*$ and 0 otherwise. For any bipartition b and tree T^*, we have $\mathbf{1}_b(T^*) \leq \text{TS}(b,T^*)$. The FBP is equal to the average of $\mathbf{1}_b(T^*)$ over the set of bootstrap trees, while the TBE is equal to the average of $\text{TS}(b,T^*)$. Thus, when using the same set of bootstrap trees, we necessarily have $\text{FBP}(b) \leq \text{TBE}(b)$. When b is a cherry ($p = 2$), we have $\mathbf{1}_b(T^*) = \text{TS}(b,T^*)$ and thus $\text{FBP}(b) = \text{TBE}(b)$. With deeper bipartitions, we generally observe that in the presence of a clear phylogenetic signal, only a small number of taxa need to be transferred to make b identical to a bipartition in T^*, while the strict presence of b in T^* can be relatively rare; the difference between $\text{FBP}(b)$ and $\text{TBE}(b)$ can then be substantial.

The transfer distance and index are related to parsimony. The branch b is equivalent to a binary $\{0,1\}$ character; assuming that the tips of T^* are labelled accordingly, we can define $\text{PA}(b,T^*)$, which is the minimum number of changes along T^* branches required to explain the labels of the tips. When b belongs to T^*, we have $\text{PA}(b,T^*) = 1$, and the more shuffled the zeros and ones among the tips of T^*, the higher is $\text{PA}(b,T^*)$. It is easy to see that $\text{PA}(b,T^*) \leq \phi(b,T^*) + 1$. Indeed, let b^* be a branch in T^* such that $\phi(b,T^*) = \delta(b,b^*)$ and assume, without loss of generality, that $\delta(b,b^*)$ is equal to the number of tips labelled 1 in the light side of b^* plus the number of tips labelled 0 in the heavy side of b^* (in other words, the light side of b^* is mostly 0 and the heavy side is mostly 1). Now consider that all internal nodes in the light side are 0 and all internal nodes in the heavy side are 1; this implies a number of changes equal to $\phi(b,T^*) + 1$, which by the definition of parsimony is larger than or equal to $\text{PA}(b,T^*)$. Parsimony is thus another option to measure branch presence, but it is inappropriate in our context. For example, consider a reference branch $b = AB|CD$, in which A, B, C and D are four large 'corner' subtrees,

and a tree T^* with an internal branch b^* grouping the corner subtrees the other way around (for example, $b^* = AC|BD$, meaning that A and C sit on one side of b^*, and B and D on the other side). Then, $\text{PA}(b,T^*)$ is equal to 2, a very low value, whereas T^* is phylogenetically very different from b because both clades defined by b are mixed. In this case, the transfer index between b and T^* is much larger and equal to the minimum size of A, B, C and D.

Recursive computation of the transfer index. A recursive algorithm to compute all transfer distances between any given bipartition b of T and all bipartitions of another tree T' has previously been described[26,29]. This algorithm is easily transformed to compute the transfer index. The principle is as follows:

(1) Map all the leaves of the light side of b to 0, the others to 1 and apply the same mapping to the leaves of T^*. Furthermore, root T^* at any internal node.

(2) With a single post-order tree traversal, one can compute the number of leaves labelled 0 and the number of leaves labelled 1 for every subtree in T^*.

(3) Let l_0 be the number of leaves labelled 0 and l_1 be the number of leaves labelled 1 in the subtree attached below a given bipartition b^*. The transfer distance between b and b^* is given by $\delta(b,b^*) = \text{Min}\{p - l_0 + l_1, l - p - l_1 + l_0\}$ (think to the missing zeros and the ones to be removed in b^* below subtree, and vice versa). This distance can be computed during the post-order traversal as well as the transfer index $\phi(b,T^*)$, which is the minimum of $\delta(b,b^*)$ for all bipartitions of T^*.

This algorithm has linear time complexity, and thus computing TBE for all bipartitions in T with r bootstrap replicates has a time complexity in $O(rl^2)$. FBP has the same time complexity, but very efficient implementations have been developed (for example, using bit vectors to encode bipartitions). In practice, computing all TBE supports with 4,000 taxa and 1,000 replicates requires less than one hour (5 core Intel Xeon 3.5 GHz), which is negligible compared to the time required to infer the reference and bootstrap trees.

Expected transfer index with random trees and TBE distribution. We have seen that the transfer index satisfies $\phi(b,T^*) \leq p - 1$. We show here that the expected transfer index is very close to this upper bound with random 'bootstrap' trees when the number of taxa is large enough. Consequently, the transfer bootstrap expectation of any branch b ($\text{TBE}(b) = 1 - \overline{\phi(b,T^*)}/(p-1)$) is close to 0 when the bootstrap trees seem to be random and do not contain any signal regarding b. This property explains why moderate supports—for example, 70% as used throughout this paper—are sufficient to reject poor branches, as a branch support of 70% cannot be observed by chance.

We first provide a simple argument to explain this result, based on the expected transfer distance between a fixed bipartition b and a random bipartition b^* with fixed light-side size p^*. Let $x = p/l$ denote the proportion of taxa in the light side of b ($x \leq 1 - x$ because $p \leq l - p$). Both bipartitions b (fixed) and b^* (random) define four taxon subsets, the sizes of which follow hypergeometric distributions with expectations: $E(\text{light side of } b \cap \text{light side of } b^*) = xp^*$; $E(\text{light side of } b \cap \text{heavy side of } b^*) = x(l - p^*)$; $E(\text{heavy side of } b \cap \text{light side of } b^*) = (1 - x)p^*$; and $E(\text{heavy side of } b \cap \text{heavy side of } b^*) = (1 - x)(l - p^*)$. It is easily seen that under these assumptions, the expected transfer distance between b and b^* is equal to the sum of the second and third (anti-diagonal) terms: that is, $E[\delta(b,b^*)] = (1 - 2x)p^* + p$. As $p^* > 0$ and $(1 - 2x) \geq 0$, we have: $E[\delta(b,b^*)] \geq p$. This result shows that the expected transfer distance between b and b^* is larger than or equal to p, for any value of p and p^*. Moreover, with a lower p^*, the expected transfer distance is closer to p. As a first approximation, we thus see that the transfer index should be close to its upper-bound $p - 1$, because it is equal to the minimum of distances which taken separately are all expected to be larger than p.

However, these distances fluctuate around their expected values, and their minimum may be lower than the minimum of their individual expectations, especially when using small samples (that is, low number of taxa). We performed computer simulations to measure the extent of this phenomenon and the validity of the $E[\phi(b,T^*)] \approx p - 1$ approximation. We used four tree sizes: $l = 16$, 128, 1,024 and 8,192 taxa, and four models of random phylogenetic trees: caterpillars (fully imbalanced), PDA, Yule–Harding and perfectly balanced[38]. For the bipartition b, all possible integer values of p in the $[2, l/2]$ range were used. The number of random bootstrap trees was equal to 1,000, and we performed 100 runs per tree size.

Results are displayed in Extended Data Fig. 1. With $l \geq 1,024$, the average transfer index with random trees is very close in relative value to the upper bound $p - 1$, and the approximation is already satisfying with $l = 128$. Moreover, the results are nearly the same for the four random tree models, suggesting that the property holds in a number of settings. As expected, the approximation is better with small p. Indeed, note that the upper bound $p - 1$ is obtained with a trivial bipartition b^* made of a unique taxon belonging to the light side of b. When a cherry in T^* contains two taxa from the light side of b, then $\phi(b,T^*) \leq p - 2$. Similar deviations are observed with subtrees in T^* containing a large fraction of taxa belonging to the light side of b. With a larger p, there is a higher probability for such an event to occur. Note, however, that large values of p (that is, $p \approx 2$) are relatively rare for most tree models (for example, Yule–Harding). Looking at the distribution of TBE, we see that having TBE larger than a moderate threshold (such as 50%)

is very unlikely, even with 16 taxa, thus explaining that TBE rarely supports poor branches with real and simulated data (Fig. 2c, d and Extended Data Figs. 2, 3, 7, 8).

Software programs and web server. We developed several tools to compute the transfer bootstrap. We first implemented a command line tool in C, 'Booster' (open source, available at https://github.com/evolbioinfo/booster). This tool computes TBE as well as FBP supports, and the stability scores of the taxa (globally or per branch). It takes two files as input: (1) a reference tree file in Newick format and (2) a bootstrap tree file in Newick format, containing all bootstrap trees. A number of software programs can be used to infer trees from multiple sequence alignments (MSAs) and produce these reference and bootstrap files in the desired format; these include RAxML, FastTree and PhyML—used in this article—as well as many others (see examples in Booster GitHub repository).

We also developed 'BoosterWeb' (http://booster.c3bi.pasteur.fr), a freely available web interface that enables users to compute bootstrap supports (TBE and FBP) easily without installing any tool on their own computer. Computations are launched on the Institut Pasteur cluster throughout a Galaxy instance. As with the command line tool, this includes the option to input reference and bootstrap trees inferred using any phylogenetic program. Another option is to upload an MSA and then run PhyML-SMS[39] (for medium-size datasets) or FastTree (for large datasets) to infer the trees. We propose a basic visualization of the resulting tree highlighting highly supported branches at a given threshold. The resulting tree can be uploaded in one-click on iTOL[40] for further manipulation. Moreover, BoosterWeb is self-contained and can be easily installed on any desktop computer (Windows, MacOS and Linux) by downloading the BoosterWeb executable.

For the sake of reproducibility, all analyses described in this article were implemented in the NextFlow workflow manager[41], and are accessible along with all our data at https://github.com/evolbioinfo/booster-workflows. The software programs that we developed to manipulate data are available for download at http://github.com/fredericlemoine/goalign and http://github.com/fredericlemoine/gotree, for manipulating alignments and trees, respectively.

Mammal dataset and analyses. We downloaded all aligned mammals COI-5P amino acid sequences from the Barcode of Life Data System (http://www.barcodinglife.org, accessed September 2015). We removed all sequences shown to be identical among several species, kept one sequence per species (several gene versions are available for some species, but no paralogues), and converted the resulting multiple alignment (1,449 sequences, 527 sites) into FASTA format. This alignment was subsampled to study the effect of tree size. We randomly drew 8 samples with 1/8th of the sequences (that is, 181) and 64 samples with 1/64th of the sequences (that is, 22). We then generated 1,000 bootstrap alignments for the full alignment and each of the 72 subsampled alignments by drawing sites with replacement.

We used FastTree[34] (options: -nopr -nosupport -wag -gamma) to infer trees from each of these reference $(1 + 8 + 64 = 73)$ and bootstrap (73,000) alignments. To ensure that the results and conclusions were independent of the tree inference method, we also performed the same analyses using RAxML with rapid bootstrap[17] (options: -f a -m PROTGAMMA -c 6 -T 10 -p $RANDOM -x $RANDOM -#1000). The FBP and TBE supports for the (73×2) reference trees were computed using Booster (command-line version written in C). All trees were drawn using iTOL and are available in the Booster GitHub repository, along with the sequence alignments. To assess whether rogue taxa removal improves FBP supports, we ran RAxML rogue-detection tool[24] (options: -J MR_DROP -z bootstrap_trees -m PROTGAMMAWAG -c 6 -T 4) and recomputed FBP supports without the detected taxa.

The FastTree and RAxML complete tree topologies were compared to the NCBI taxonomy (https://www.ncbi.nlm.nih.gov/taxonomy), which was converted to Newick format and reduced to the 1,444 taxa common to both our alignment and the NCBI taxonomy. This NCBI tree is not fully resolved and summarizes common belief about the evolutionary history of mammals, resulting from a number of phylogenetic studies based on numerous markers. The unresolved part of the NCBI tree (~4.35 descendants per node on average, instead of 2 for a fully resolved tree) corresponds to the unknown or uncertain part of that history. To cope with uncertainty, we used quartets to compare the (fully resolved) inferred trees to the NCBI tree. A quartet is a tree topology with four taxa; $AB|CD$ is the standard notation for quartets, indicating that taxa A and B form a cherry separated by an internal branch from the cherry formed by C and D; a quartet is unresolved when the four taxa are connected to a single central node. A bipartition b induces a quartet $AB|CD$ when A and B belong to the same side of b, and C and D to the other side. We used tqDist[42] to count the number of quartets induced by the reference branches, which appeared to contradict the quartets induced by the NCBI tree and its bipartitions; for example, $AB|CD$ was found in the studied branch, whereas $AC|BD$ was found in the NCBI tree. Unresolved quartets of the NCBI tree were not counted as contradictory, as they represent an unknown evolutionary truth and the inferred resolution could be correct. Such an approach would be difficult to implement with the transfer distance. The number of contradicted quartets was

divided by the total number of quartets induced by the studied branch, to obtain a normalized measurement in the [0,1] range (0: no contradiction; 1: all induced quartets are contradicted). We used the same approach to check the accuracy of the FastTree and RAxML tree topologies, comparing the whole set of quartets induced by the inferred tree to those induced by the NCBI tree.

HIV dataset and analyses. From the HIV database (https://www.hiv.lanl.gov/content/index) we retrieved *pol* sequences of the nine 'pure' subtypes of HIV-1 group M, corresponding to positions 2258–3300 relative to the HXB2 reference strain (accessed September 2014). The 'one sequence per patient' option was used and we randomly selected samples of the over-sampled subtypes (A1, B, C, D and G), resulting in a final dataset of 9,147 sequences. These sequences are annotated as 'pure' (that is, non-recombinant) in the database, using a fast filtering approach. However, 48 recombinant sequences were still detected using the standalone version of jpHMM[36] (version March 2015; options: -v HIV, with default input and priors). These 48 sequences were kept in the analyses to study the effect of recombinant sequences, as their presence is inevitable in any HIV dataset. jpHMM was also used to annotate the whole set of sequences depending on their subtype or recombinant status.

Sequences were aligned using MAFFT[43] (version 7.0; default parameters) along with the HXB2 reference strain. Codon positions associated with major drug resistance mutations were removed before tree inference, resulting in an alignment of 1,043 DNA sites (R source code available at https://github.com/olli0601/big.phylo). This alignment was subsampled to study the effect of tree size. We randomly drew 16 samples with 1/16th of the sequences (that is, 571), and 256 samples with 1/256th of the sequences (that is, 35). Then, we generated 1,000 bootstrap alignments for the full alignment and each of the 272 subsampled alignments, by drawing sites with replacement.

We used FastTree[34] (options: -nopr -nosupport -gtr -nt -gamma) to infer trees from each of these reference $(1 + 16 + 256 = 273)$ and bootstrap (273,000) alignments. The FBP and TBE supports for the 273 reference trees were computed using Booster (command-line version written in C). All trees were drawn using iTOL (http://itol.embl.de/) and are available on the Booster-workflows GitHub repository, along with the sequence alignments. The instability score was computed considering the reference branches with TBE > 70% (the signal becomes noisy when incorporating branches with lower supports in the calculation, as these branches may be erroneous and thus non-informative about taxon stability). For every taxon, the instability score is equal to the average number of times it has to be transferred to recover these branches from the bootstrap trees, divided by the number of these branches.

The most representative clades for each of the subtypes in the reference trees (Fig. 1 and Extended Data Fig. 5) were obtained by minimizing the transfer distance. For example, in Fig. 1 with the full dataset, we obtained a clade very close to subtype B, with 3,559 taxa, 2 wrong taxa (that is, non-B), and all (3,557 taxa) B taxa included, resulting in the values 3,559, 2 wrong (w) and 0 missing (m) shown in this figure.

A similar approach was used for the regional variants of subtype C, which is responsible for approximately 50% of the HIV-1 infections in the world. Three monophyletic variants of subtype C have been identified by phylogenetic analysis in East Africa[44], South America[45] and India[46]. Furthermore, the South American epidemic was shown to originate in the East African cluster[44]. To identify these variants in the inferred trees (Fig. 1 and Extended Data Fig. 5) we again used the transfer distance. Following previous publications[44–46], we extracted three groups of C sequences from the whole dataset, based on their geographic origins: East Africa (EA: 440 sequences, originating from Burundi (288), Djibouti (1), Ethiopia (9), Kenya (41), Somalia (1), Sudan (11), Tanzania (78) and Uganda (11)), India (IND: 154 sequences, originating from India (133), Nepal (13) and Myanmar (8)) and South America (SA: 14 sequences, originating from Brazil (12), Uruguay (1) and Argentina (1)). We then searched for the tree clades that were closer to these three sets of sequences. The South American sequences were not accounted for in transfer distance computations when searching for the East African clade, as they originate from East Africa. Moreover, we checked that no neighbouring, nearly optimal clade was supported by FBP. In all three cases, we found clades closely related to the sequence sets. As expected, the South American clade was included in the East African clade. The features of these clades are displayed in Fig. 1 and Extended Data Fig. 5. The fractions correspond to the number of studied sequences included in these clades, versus the total number of such sequences in the whole dataset (for example, 360 East African sequences in the East African clade in Fig. 1, among 440 in the whole tree). The 'wrong' sequences were expected in most cases. For example, the Indian clade (167 sequences, 143 from IND among 154 in the whole tree) contains 19 sequences from China corresponding to the spread of the virus in Asia via heroin trafficking routes[46].

Simulated data and analyses. The aim of our simulation experiments was to check that the results observed with the mammal and HIV-1 datasets are reproducible and quantifiable when the simulation conditions and correct tree are known,

notably regarding the support of poor branches and the ability to detect rogue taxa. Simulated data mimicked the mammal dataset. We used the tree inferred by PhyML[47] (options: -b 0 -m WAG -a e -t e -o tlr -d aa) from the full COI-5P protein alignment with 1,449 taxa. Protein sequences were evolved along this tree using INDELible[48], which was launched with options and parameter values derived from the PhyML analysis, and similar to previously conducted experiments[24] to assess the accuracy of rogue-taxon detection. The length of the root sequence was 250 AAs; the substitution model was WAG; amino acid frequencies were estimated from the COI-5P alignment; the rates across-sites model used 4 gamma categories with 'alpha' = 0.441 and no invariant sites; and the indel model used 'power law', 'parameter' = 1.5, 'indel max size' = 5, and 'indel rate' = 0.02.

In this manner, we obtained a first 'non-noisy' MSA of length ~500 with ~50% gaps. Noise was added to this MSA to mimic rogue taxa and homoplasy. We shuffled the amino acids vertically for 50% of the sites (MSA columns), thus making these sites homoplasic. For 5% of the sequences (MSA rows), 25% additional sites were shuffled vertically, thus making these sequences unstable and 'rogue', as they contained half of the phylogenetic signal compared to the other (95%) sequences. Both noisy and non-noisy MSAs were used to compare FBP and TBE. To measure the effect of tree size, both MSAs (comprising 1,449 sequences) were sampled to obtain 8 MSAs with 181 sequences (~1/8 of the full sequence set) and 64 MSAs with 22 sequences (~1/64 of the full sequence set). For each of these reference MSAs we sampled with replacement 1,000 pseudo-alignments to compare the two bootstrap methods. All trees were inferred using FastTree (options: -nopr -nosupport -wag -gamma). Just as with the mammal dataset, for each of the branches in the reference trees we computed the percentage of quartet-based conflicts with the correct (PhyML) tree used to generate the data. We also computed the instability score of all taxa in the complete noisy MSA, using only the branches with TBE > 70%.

To check the repeatability of FBP and TBE, we generated 1,000 noisy MSAs using the same phylogenetic tree, simulation procedure and set of rogue taxa as the reference noisy MSA (1,449 sequences, ~500 sites and ~50% gaps). We then compared the branch supports of the inferred branches computed using the pseudo-alignments to those obtained using the simulated MSAs. The bootstrap theory[2] indicates that both types of supports are close when the sample size is large enough. The goal was thus to check that 500 sites are enough to obtain a good approximation, and that the bootstrap-based and simulation-based supports are clearly correlated (Pearson's and Spearman's coefficients), as well as the instability score (again computed using branches with TBE > 70%). This experiment was performed with FBP and TBE, with both FastTree (options: -nopr -nosupport -wag -gamma) and RAxML (options: -f d -m PROTGAMMAWAG -c 6).

Lastly, the same experiment was used to check the validity of the plug-in principle: we compared the FBP and TBE supports of every inferred branch (both FastTree and RAxML) to the presence or absence (1/0) of that branch in the true tree (FBP), and the normalized transfer distance between that branch and the true tree (TBE).

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Code availability. Our web interface and software programs are available from Booster website (http://booster.c3bi.pasteur.fr) and GitHub (https://github.com/evolbioinfo/booster). The transfer bootstrap is available in several phylogenetic programs, including PhyML, SeaView, RAxML-NG and others (see http://booster.c3bi.pasteur.fr).

Data availability. All our multiple alignments, phylogenetic trees and workflows are available as Source Data. This material is also available from Booster website (http://booster.c3bi.pasteur.fr). All other data are available from the corresponding author upon reasonable request.

37. Robinson, D. F. & Foulds, L. R. Comparison of phylogenetic trees. *Math. Biosci.* **53**, 131–147 (1981).
38. Semple, C. & Steel, M. A. *Phylogenetics* (Oxford Univ. Press, Oxford, 2003).
39. Lefort, V., Longueville, J. E. & Gascuel, O. SMS: smart model selection in PhyML. *Mol. Biol. Evol.* **34**, 2422–2424 (2017).
40. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
41. Di Tommaso, P. et al. Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* **35**, 316–319 (2017).
42. Sand, A. et al. tqDist: a library for computing the quartet and triplet distances between binary or general trees. *Bioinformatics* **30**, 2079–2080 (2014).
43. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
44. Delatorre, E. O. & Bello, G. Phylodynamics of HIV-1 subtype C epidemic in east Africa. *PLoS ONE* **7**, e41904 (2012).
45. Soares, M. A. et al. A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. *AIDS* **17**, 11–21 (2003).
46. Siddappa, N. B. et al. Identification of subtype C human immunodeficiency virus type 1 by subtype-specific PCR and its use in the characterization of viruses circulating in the southern parts of India. *J. Clin. Microbiol.* **42**, 2742–2751 (2004).
47. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
48. Fletcher, W. & Yang, Z. INDELible: a flexible simulator of biological sequence evolution. *Mol. Biol. Evol.* **26**, 1879–1888 (2009).
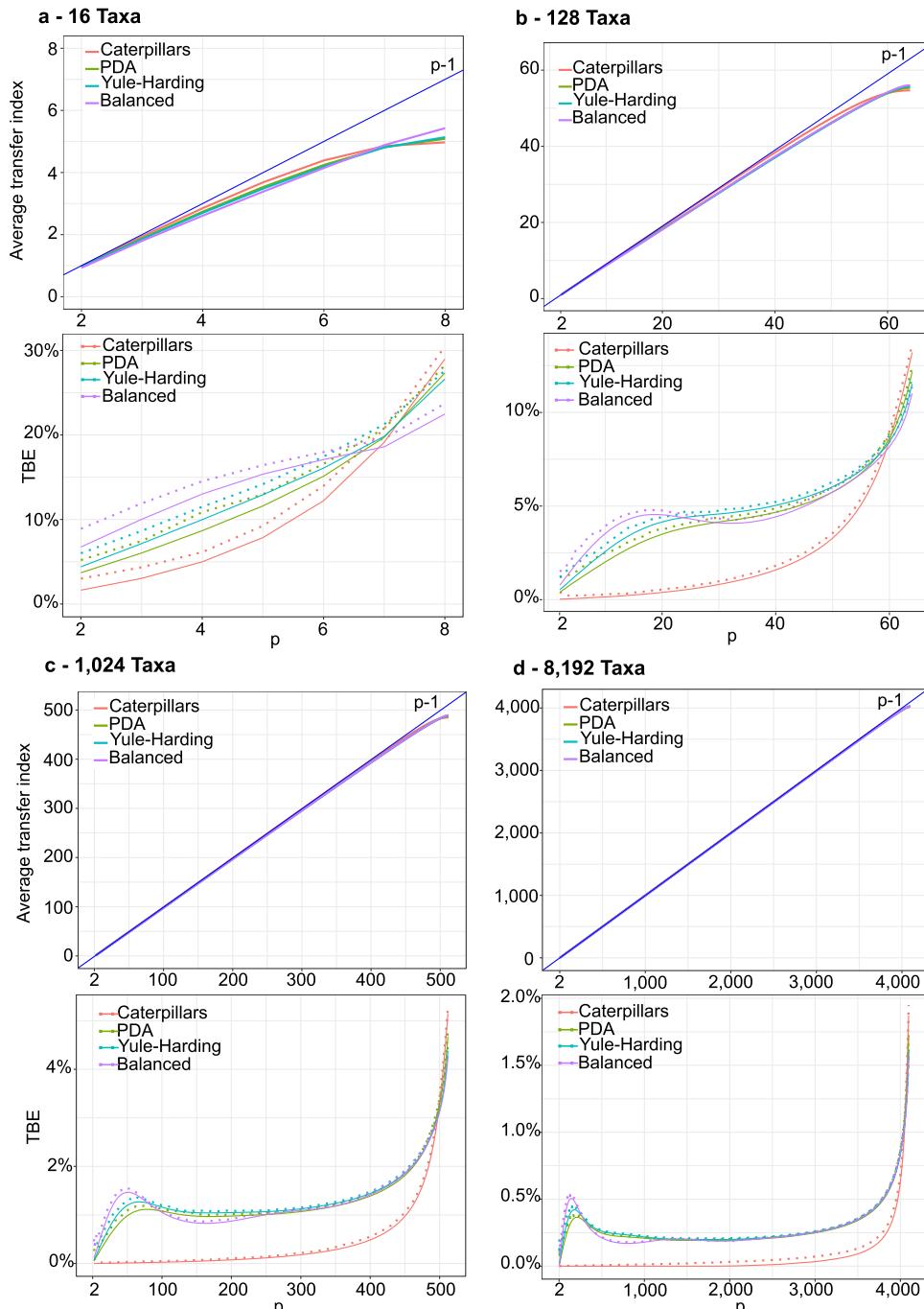
Extended Data Fig. 1 | Transfer index expectation and TBE support with random trees. a–d, For each number of taxa (16, 128, 1024 and 8, 191 in **a**, **b**, **c** and **d**, respectively) and random tree model, we compare the transfer index average over 100 runs with the upper-bound $p - 1$ (top graphs in each panel). We also compare the average transfer bootstrap support (TBE) to 0, and provide the maximum value observed among 100 runs (dashed lines), thus approximating the 1% quantile of the distribution (bottom graphs). In these experiments, the number of random 'bootstrap' trees is equal to 1,000. With $l \geq 1,024$ (**c**), the average transfer index with random trees is very close in relative value to the upper-bound $p - 1$ and the approximation is already satisfying with $l = 128$ (**b**). Furthermore, the results are nearly the same for the four random tree models, suggesting that the asymptotic behaviour holds in a number of settings. As expected, the approximation of the transfer index over random bootstrap trees by $p - 1$ is better with small values of p. These results explain why moderate TBE supports—for example, 70% as used in this article—are sufficient to reject poor branches, as a TBE branch support of 70% cannot be observed by chance, even with a small number of taxa (for example, 16, as in **a**).
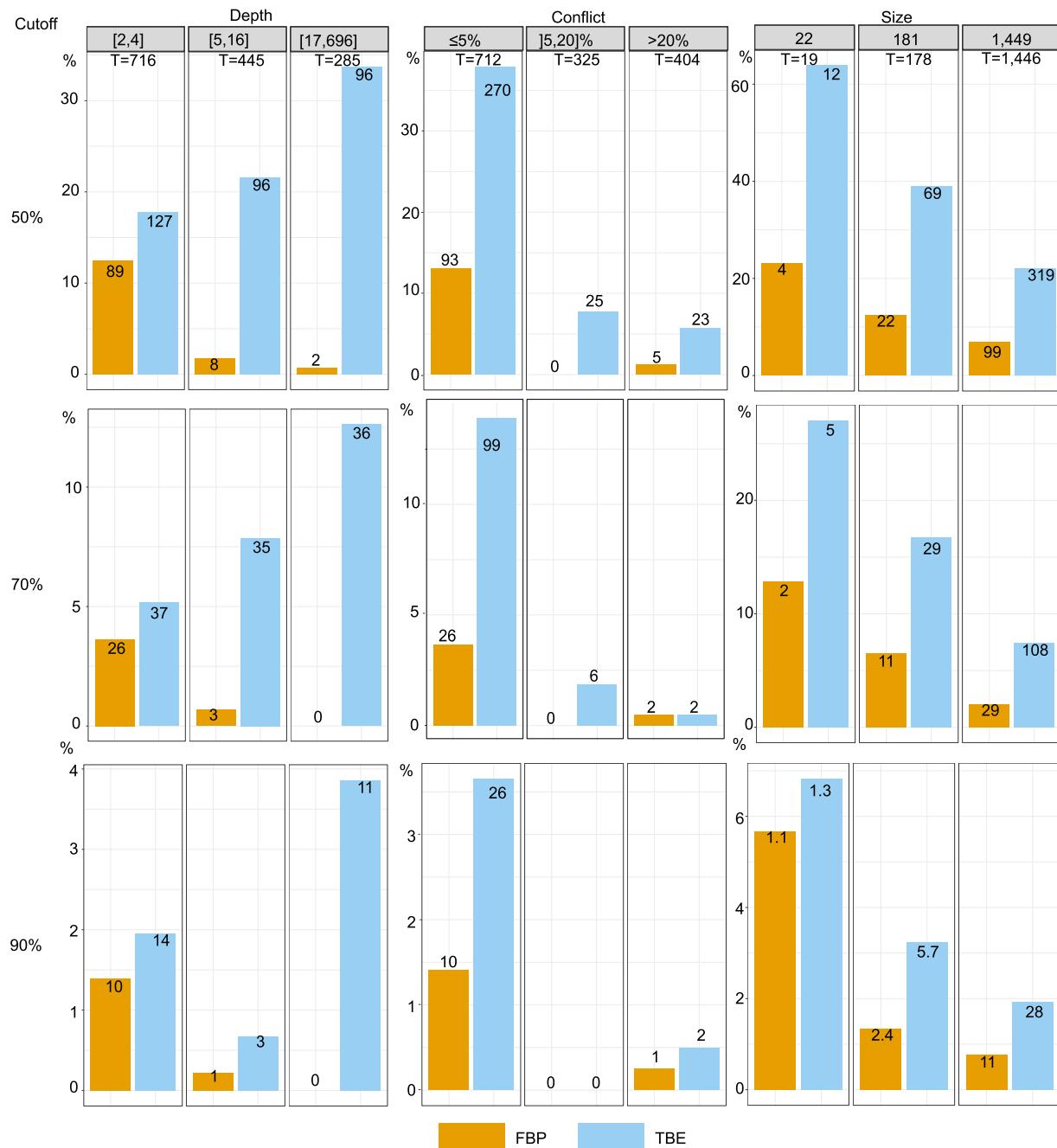
**Extended Data Fig. 2 | Comparison of FBP and TBE using the mammal dataset and FastTree phylogeny.** FBP and TBE supports are compared with respect to branch depth, quartet conflicts with the NCBI taxonomy and tree size (see main text and legends of Figs. 1, 2 for explanations). Three support cut-offs are used to select the branches: 50%, 70% and 90% (for example, 28 branches among the 1,446 in total have TBE ≥ 90% and 11 have FBP ≥ 90%). The FastTree topology is poor, with 38% of quartets contradicted by the NCBI taxonomy, and 404 of the 1,441 branches with contradictions above 20%. Despite this difficulty, FBP and TBE perform well: they give supports larger than 70% to a very low number of moderately ((5,20]%) and highly (> 20%) conflictual branches. FBP supports very few deep branches, whereas TBE supports a larger number of branches and is especially useful with large trees. Comparing the three cut-offs, we see that with a 50% cut-off the selected branches are still weakly contradicted, especially with FBP; as expected, with TBE the fraction of contradicted branches (> 5%) is a bit higher but still low (~7%). With a cut-off of 90% very few branches are selected (~2% with TBE), thus justifying the use of the 70% threshold for TBE—as is standard with FBP.
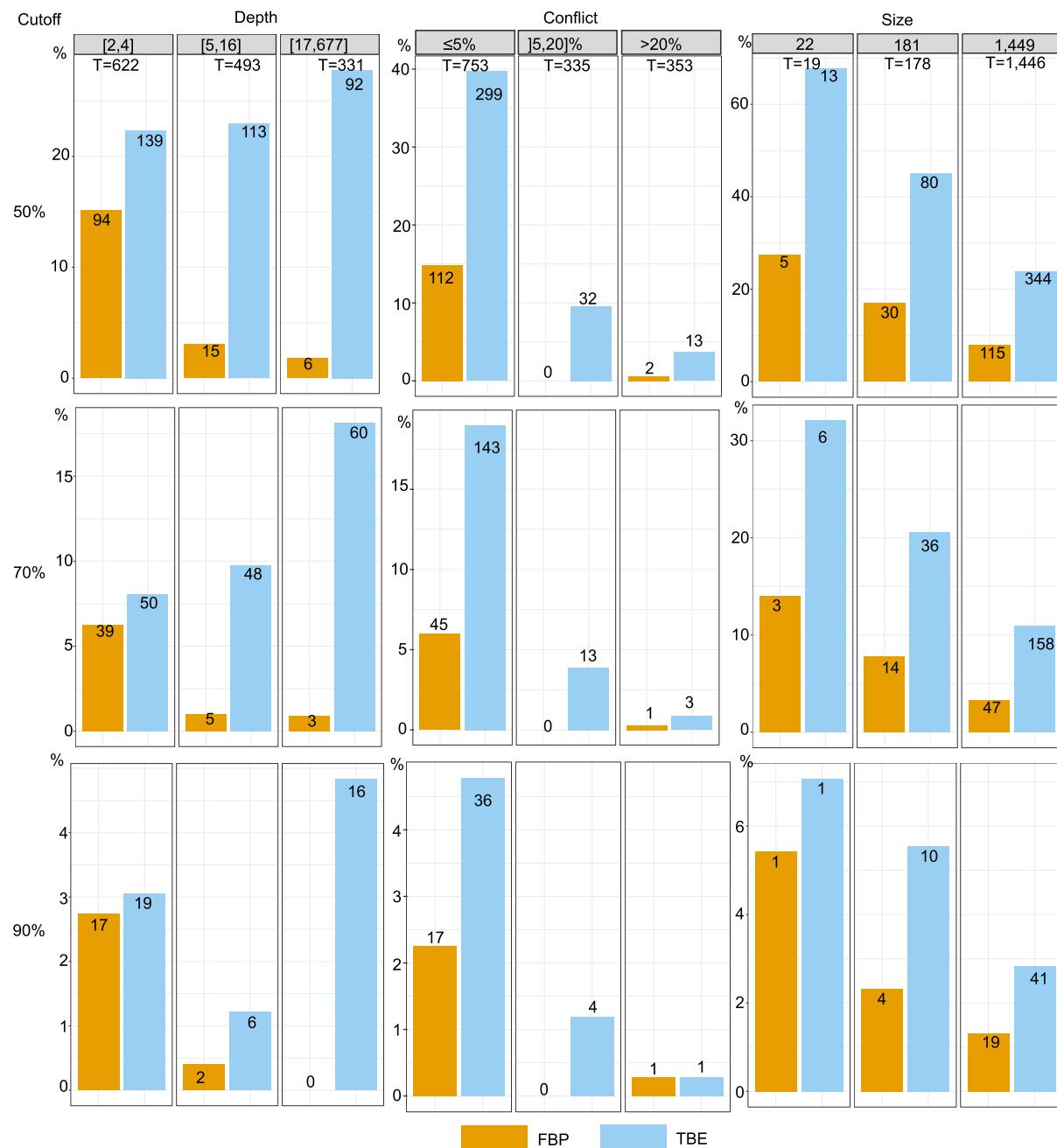
Extended Data Fig. 3 | Comparison of FBP and TBE using the mammal dataset and the phylogeny inferred by RAxML with rapid bootstrap. FBP and TBE supports are compared with respect to branch depth, quartet conflicts with the NCBI taxonomy and tree size (see main text and legends of Figs. 1, 2 for explanations). Three support cut-offs are used to select the branches: 50%, 70% and 90% (for example, 41 branches among the 1,446 in total have TBE $\geq$ 90% and 19 have FBP $\geq$ 90%). The RAxML topology is closer to the NCBI taxonomy than is the FastTree topology (27% versus 38% of contradicted quartets, and 353 versus 404 branches with contradiction > 20%, respectively). However, the RAxML topology is still relatively poor, as expected in this type of phylogenetic study based on a unique marker (Fig. 4 and main text). Despite this difficulty, FBP and TBE perform well as they give supports larger than 70% to a very low number of moderately ((5,20]%) and highly (> 20%) conflictual branches. The supports obtained with RAxML are higher than those obtained with FastTree (47 versus 29 branches with FBP > 70% for RAxML and FastTree,

respectively; 158 versus 108 branches with TBE > 70% for RAxML and FastTree, respectively). Part of the explanation could be that the RAxML tree is more accurate than that of FastTree, and is thus better supported. Another factor is that the rapid bootstrap tends to be more supportive than the standard procedure, as shown in previous publications[16]. Indeed, the rapid bootstrap uses already inferred trees to initiate tree searching, and therefore tends to produce less diverse bootstrap trees than the standard, slower procedure, which restarts tree searching from the very beginning for each replicate. Despite these differences between FastTree and RAxML with rapid bootstrap, similar conclusions are drawn when comparing FBP and TBE: FBP supports very few deep branches, whereas TBE supports a larger number of them; TBE is especially useful with large trees; and both methods support a very low number of contradicted branches. Comparing the support cut-offs, 70% again appears as a good compromise for both FBP and TBE.
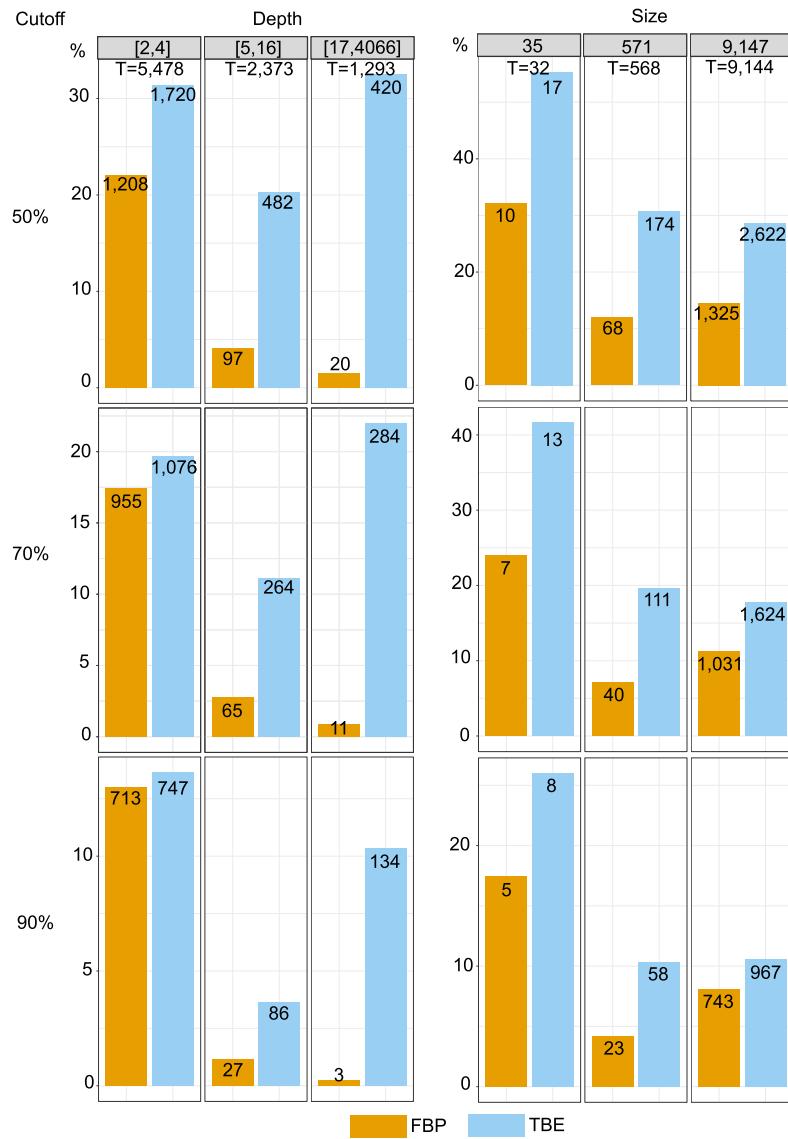
Extended Data Fig. 4 | Comparison of FBP and TBE using the HIV dataset and FastTree phylogeny. FBP and TBE supports are compared with respect to branch depth, and tree size (see main text and legends of Figs. 1, 2 for explanations). Three support cut-offs are used to select the branches: 50%, 70% and 90% (for example, 1,624 branches among the 9,144 in total have TBE > 70% and 1,031 have FBP > 70%). Results are for the most part similar to those observed with the mammal dataset. We see a major effect of depth on FBP supports: with the full dataset, less than 1% of the deep ($p > 16$) branches have FBP support larger than 70%, whereas this percentage is higher than 20% with TBE. The effect of tree size is less pronounced. The fraction of supported branches decreases when the tree size increases from 35 to 571 taxa, but is analogous between 571 and 9,147 taxa. Furthermore, the gap between FBP and TBE remains similar, probably owing to the very large number of cherries and small clades, for which TBE and FBP are nearly equivalent. Regarding the support cut-off, 70% again appears as a good compromise for TBE, though there is no way to evaluate the fraction of supported branches that is actually erroneous. The interpretability of TBE will be a major asset for choosing the support level depending on the phylogenetic question being addressed. Here, as recombinant sequences are inevitable, lower supports than with mammals are likely to be acceptable.

a - Subsample 1 - FBP & TBE > 70%
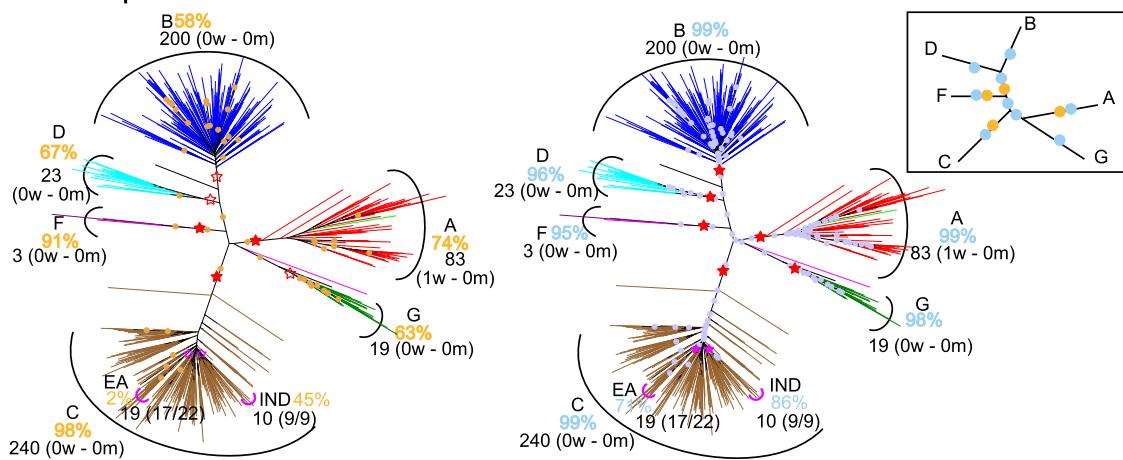
b - Subsample 2 - FBP & TBE > 70%
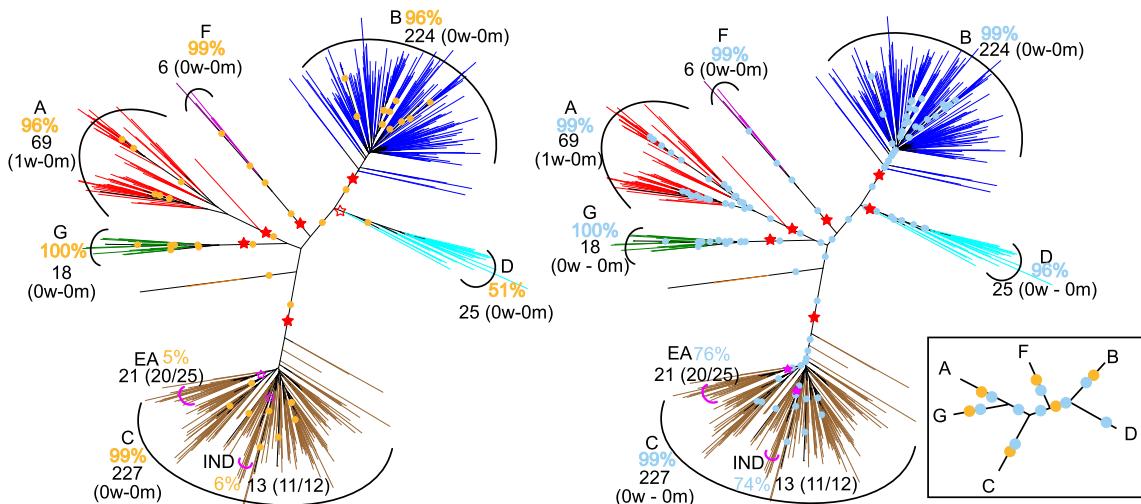
c - Deep branching - Full dataset supports
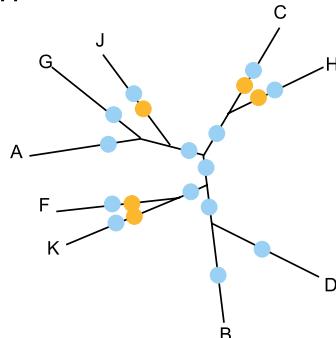
Extended Data Fig. 5 | Subtype deep branching and comparison of FBP and TBE using medium-sized HIV datasets. As the taxa were randomly drawn from the full dataset, the supports and findings show some fluctuations. **a**, **b**, Trees obtained with two of the medium-sized datasets; branches with FBP > 70%: yellow dots; branches with TBE > 70%: blue dots; subtype clades: red stars, filled if support > 70% (see Methods and Fig. 1 legend for further details). **c**, Deep branching of the subtypes[19] and supports obtained on the full dataset (see also Fig. 1). Rare subtypes (H, J and K) are absent in the medium-sized datasets, and the subtype clades are almost perfectly recovered (only one incorrect taxon in A

clade for both trees). FBP supports are higher when using medium-sized datasets than when using the full dataset (for example, 58% and 99% for subtype B, versus 3% in Fig. 1). However, some subtype clades (for example, D) have moderate FBP support, though the clade matches the subtype perfectly. When using TBE, all subtype supports are higher than 95%. The deep branching is the same for all full and medium-sized datasets, and is identical to that found in a previous study[19], but is not supported by FBP, whereas TBE is larger than 70% for every branch (or path in Fig. 1). Again, the Indian and East African sub-epidemics of subtype C are supported by TBE, but not by FBP.
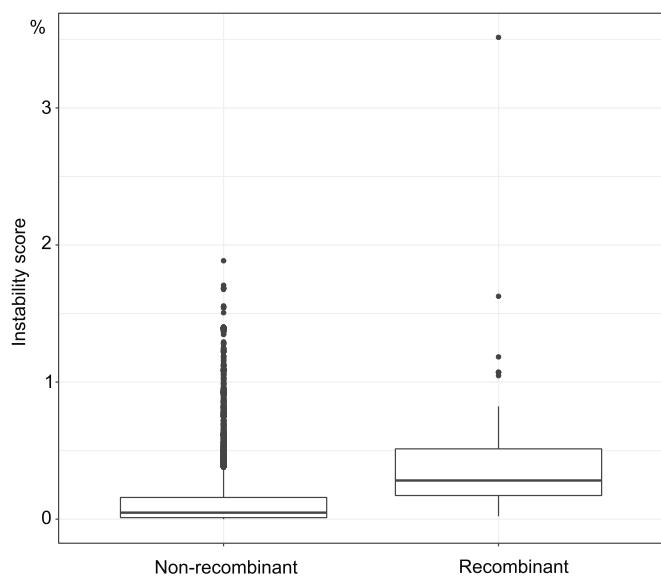
Extended Data Fig. 6 | Distribution of the instability score in HIV recombinants. We see a clear difference between the distributions of the instability score for the recombinant and non-recombinant sequences, which means that this score can be used to detect or confirm the recombinant status of sequences (box quantiles: 25%, 50% and 75%). See main text for details.
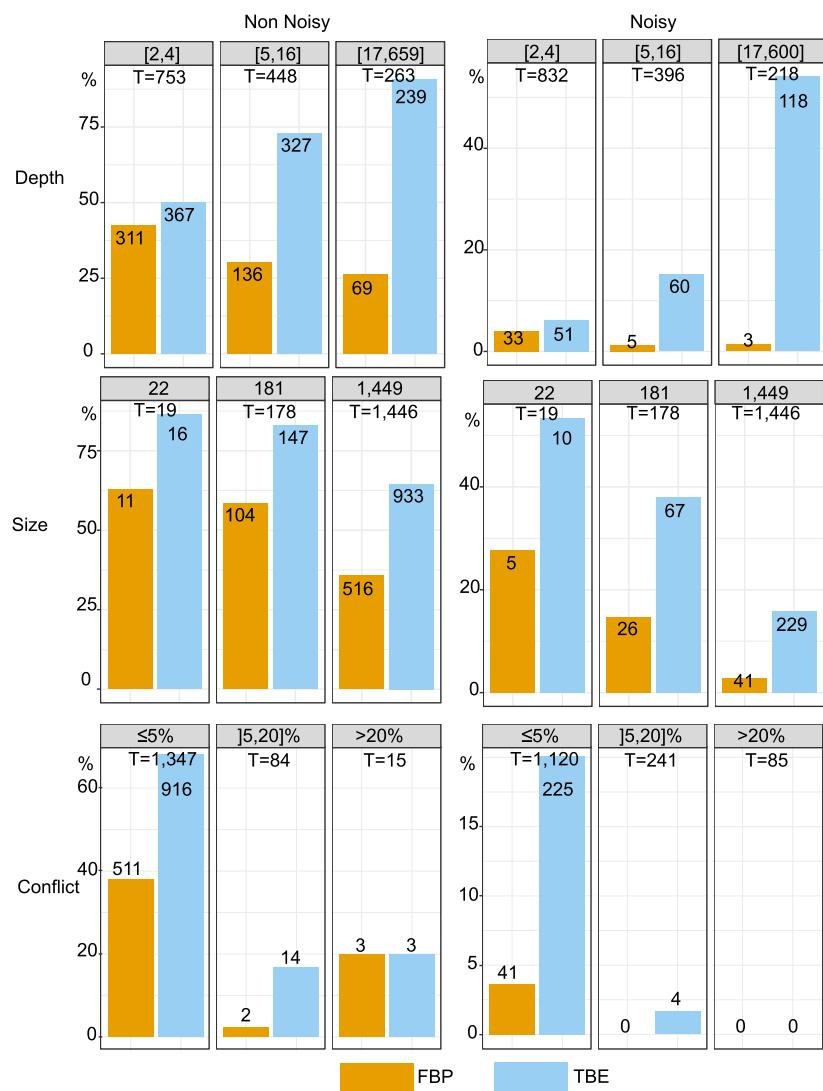
Extended Data Fig. 7 | Comparison of FBP and TBE using non-noisy and noisy simulated data. Noisy data include rogue taxa and homoplasy and non-noisy data do not (see Methods for details). The graphs display the distribution of branches with FBP or TBE support > 70%. Supports are compared regarding branch depth, tree size and quartet conflicts with the model tree used for simulations (see main text and legends of Figs. 1, 2 for explanations). Results are fully congruent with those obtained with real datasets. TBE supports more deep branches than FBP, especially with noisy data. The effect of tree size is also more visible with noisy MSA, and the number of supported branches with moderate ((5,20]%) and high (> 20%) conflict levels is very low, for both FBP and TBE.
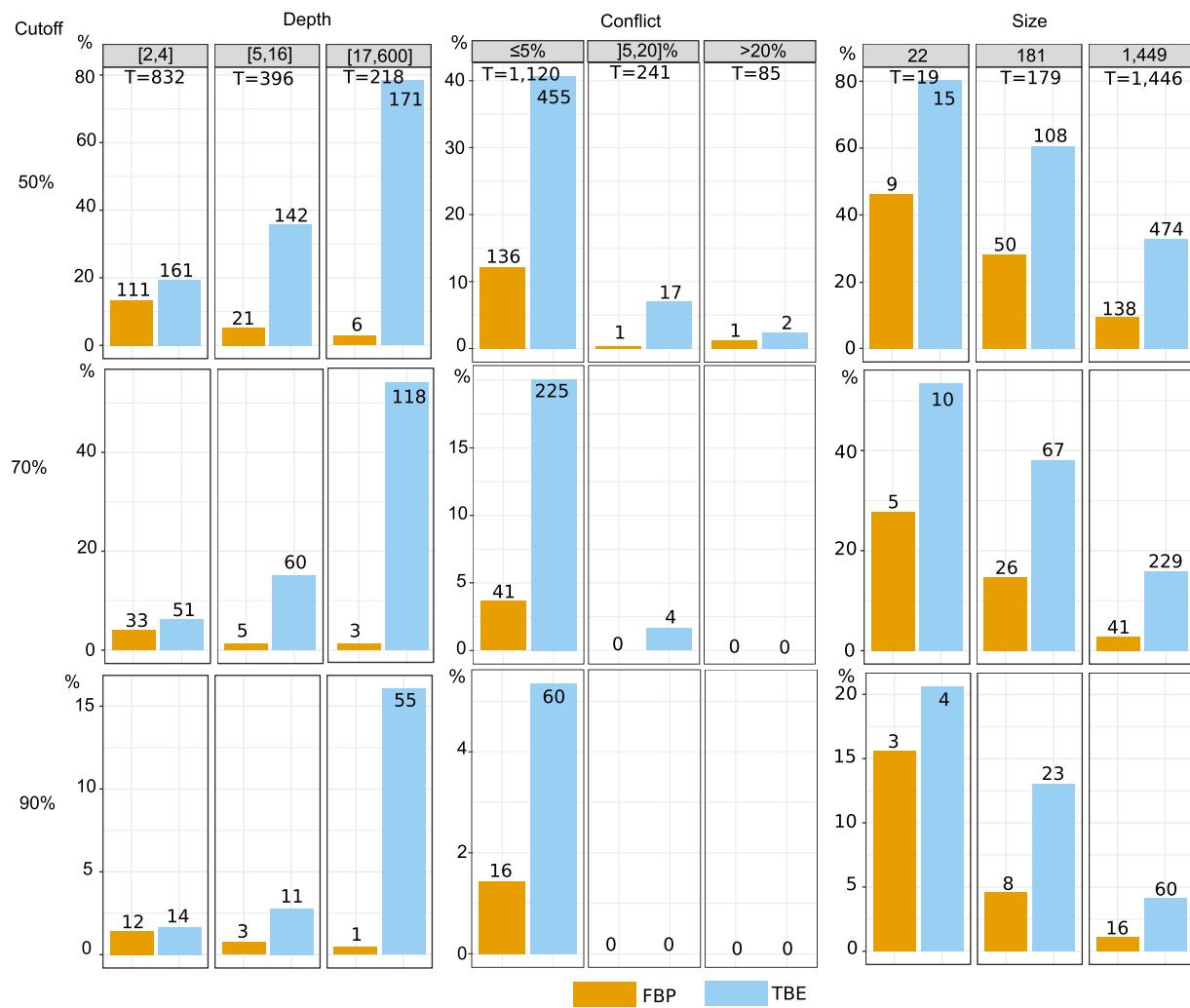
Extended Data Fig. 8 | Comparison of FBP and TBE at different support cut-offs using simulated, noisy data. Comparison of FBP and TBE with respect to branch depth, quartet conflicts and tree size, at different support cut-offs (see main text and legends of Figs. 1, 2 for explanations). A cut-off of 50% seems to be acceptable, as neither FBP nor TBE support highly contradicted branches. However, this could be due to the low level of contradiction compared to real datasets (85 branches with contradiction > 20%, versus about 400 in the mammal dataset in Extended Data Figs. 2, 3).
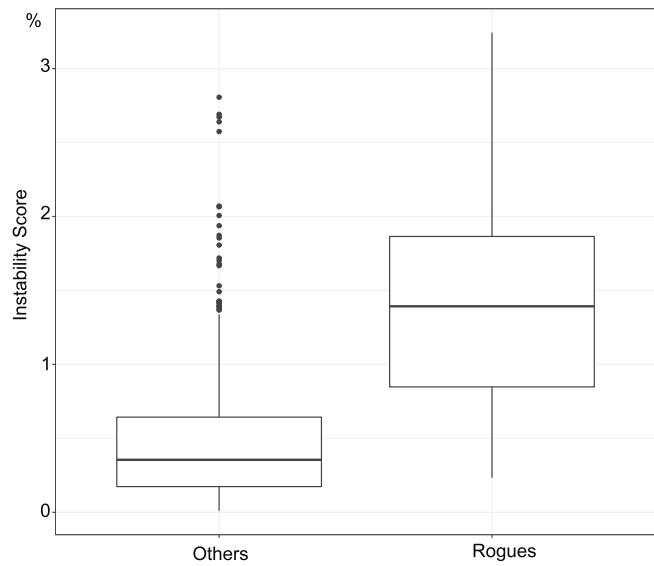
Extended Data Fig. 9 | Distribution of the instability score in rogue taxa using simulated, noisy data. TBE again appears to be useful for detecting and confirming rogue taxa (box quantiles: 25%, 50% and 75%). See main text for details.

a - FBP Boot vs. Simu

Branch supports: bootstrap trees (y-axis)
Branch supports: simulation-based trees (x-axis)
$\rho = 0.75$

b - TBE Boot vs. Simu

Branch supports: bootstrap trees (y-axis)
Branch supports: simulation-based trees (x-axis)
$\rho = 0.85$

c - TBE Boot vs. True

Branch supports: bootstrap trees (y-axis)
Inferred tree vs. True tree (x-axis)
$\rho = 0.74$

d - TBE unstable taxa

Rogue fraction among x most unstable taxa (y-axis)
x most unstable taxa (x-axis)
TPR=58%, FPR=4%
TPR=48%, FPR=5%
Bootstrap trees
Simulation-based trees

e - Correlations: Pearson(Spearman)

	%Correct	FBP Simu/Boot	FBP True/Boot	TBE Simu/Boot	TBE True/Boot	%Rogues
RAxML	322/1,446	0.75(0.83)	0.59(0.54)	0.85(0.83)	0.74(0.70)	58
FastTree	261/1,446	0.72(0.82)	0.54(0.48)	0.85(0.82)	0.75(0.72)	60

Extended Data Fig. 10 | Repeatability and accuracy of FBP and TBE using simulated data. The bootstrap theory[1,2] indicates that with large samples the supports estimated using bootstrap replicates should be close to supports obtained with datasets of the same size drawn from the same distribution as the original sample. We used simulated data to check that this property holds with protein MSAs of 1,449 taxa and about 500 sites (see main text for details). **a, b,** Comparison of the two supports (**a,** FBP; **b,** TBE) for all branches in the tree inferred by RAxML from the original MSA. We observe a clear correlation, which is higher for TBE ($\rho = 0.85$) than for FBP ($\rho = 0.75$) using Pearson's linear correlation coefficient, but identical (0.83) using Spearman's rank coefficient, which is better suited to the discontinuous nature of FBP. These results appear to contradict previous conclusions[7] that the bootstrap is a highly imprecise measure of repeatability. However, this previous work measured the probability of inferring the correct tree (not the supports of inferred branches, as consistent in the bootstrap context) and its main result was based on 50 sites, which is probably too low for the bootstrap theory to apply.

The bootstrap also relies on the plug-in principle[2,3,6,9], which states that the distribution of the distance between the true tree and the inferred tree can be well-approximated by the distribution of the distance between the inferred and bootstrap trees. **c,** The accuracy of TBE in predicting the topological distance between b and the true tree as measured using the normalized transfer index, for every branch b inferred by RAxML from the original MSA. Again, we observe a clear correlation ($\rho = 0.74$, Spearman's rank coefficient = 0.70). We performed the same experiment with FBP, seeking to predict the presence or absence (1/0) of the inferred branch in the tree true; a lower but significant correlation was found ($\rho = 0.59$, Spearman's rank coefficient = 0.54). **d,** Comparison using RAxML of the performance of simulation-based and bootstrap-based instability scores in detecting rogue taxa; both are nearly identical. TPR, true positive rate; FPR, false positive rate. **e,** Table summarizing the results described above and those of FastTree, which are nearly identical to those of RAxML, except regarding topological accuracy (%Correct, fraction of correct branches) for which RAxML is again more accurate than FastTree.

ARTICLE

Intra-tumour diversification in colorectal cancer at the single-cell level

Sophie F. Roerink[1,13], Nobuo Sasaki[2,11,13], Henry Lee-Six[1,13], Matthew D. Young[1], Ludmil B. Alexandrov[3,4,5], Sam Behjati[1,6], Thomas J. Mitchell[1,7], Sebastian Grossmann[1], Howard Lightfoot[1], David A. Egan[8,12], Apollo Pronk[9], Niels Smakman[9], Joost van Gorp[10], Elizabeth Anderson[1], Stephen J. Gamble[1], Chris Alder[1], Marc van de Wetering[2], Peter J. Campbell[1], Michael R. Stratton[1]* & Hans Clevers[2]*

Every cancer originates from a single cell. During expansion of the neoplastic cell population, individual cells acquire genetic and phenotypic differences from each other. Here, to investigate the nature and extent of intra-tumour diversification, we characterized organoids derived from multiple single cells from three colorectal cancers as well as from adjacent normal intestinal crypts. Colorectal cancer cells showed extensive mutational diversification and carried several times more somatic mutations than normal colorectal cells. Most mutations were acquired during the final dominant clonal expansion of the cancer and resulted from mutational processes that are absent from normal colorectal cells. Intra-tumour diversification of DNA methylation and transcriptome states also occurred; these alterations were cell-autonomous, stable, and followed the phylogenetic tree of each cancer. There were marked differences in responses to anticancer drugs between even closely related cells of the same tumour. The results indicate that colorectal cancer cells experience substantial increases in somatic mutation rate compared to normal colorectal cells, and that genetic diversification of each cancer is accompanied by pervasive, stable and inherited differences in the biological states of individual cancer cells.

Recent studies have explored genetic diversification within cancer cell populations by identifying mutations shared by subpopulations of cells within the cancer clone[1–6]. In principle, however, the extent of intra-tumour genetic diversity is most comprehensively revealed by single-cancer-cell DNA sequencing, which can potentially identify all mutations, including those that arose in the distant past, those that occurred very recently, and those not present in any other cell[7,8]. Despite recent advances, however, single-cell genome sequencing remains dependent on prior whole-genome amplification, which is associated with incomplete genome coverage and artefactual mutations[9]. Diversification of epigenomic, transcriptomic, proteomic and metabolic states, and of functional states such as resistance to anticancer therapy, may also occur during expansion of the neoplastic cell population[10–13]. Methylation, gene expression and drug response data have previously been obtained from multiple cells from individual tumours, but the collection of all these features together with accurate genome information from the same single cells has not been reported, to our knowledge[14–19]. The origins of epigenomic and transcriptomic diversification are unclear, and there is little insight into whether these are transient or stable.

One experimental approach that enables comprehensive, systematic and integrated exploration of intra-tumour diversification is to derive immortal cell lines from multiple single cells from the same cancer[20]. These serve as proxies for the single cells from which they originate and can be subjected to extensive and multimodal characterization, thereby revealing all aspects of intra-tumour diversification retained during in vitro growth. We have recently developed protocols for derivation of clonal organoids from normal and neoplastic colorectal stem cells, and we use this strategy here to compare single cells from normal and neoplastic colorectal epithelium[21–24].

Comparison of the number of mutations in single cancer cell genomes with that in individual normal cells from the same tissue may also reveal whether alterations in somatic mutation rate and mutational process have been experienced by neoplastic cells. Substantial increases in mutation rate are known to occur during the development of cancers with DNA mismatch repair deficiency or mutations in genes encoding DNA polymerases[25]. Whether mutation rate increases are common in cancers without these specific abnormalities is, however, currently unknown and a matter of controversy[26–30].

Clonal organoid derivation

Colorectal cancers from three previously untreated individuals (P1, P2 and P3) were each dissected into 4–6 pieces (Extended Data Fig. 1). Organoid cultures were derived from cell suspensions made separately from each piece and were maintained for up to one week without passage. Subsequently, individual organoids were disaggregated and flow-sorted to obtain single cells from which clonal cancer organoids were established. For each individual, organoids were also derived from single crypts in normal colorectal epithelium from the same resection specimens. A crypt derives from a single stem cell that has been estimated to exist several months before crypt isolation[24,31].

The coding regions of 360 known cancer genes were sequenced in all normal and cancer-derived clonal organoids for likely driver mutations and a subset were whole-genome sequenced. Somatic mutations

[1]Cancer Ageing and Somatic Mutations Programme, Wellcome Trust Sanger Insitute, Hinxton, UK. [2]Hubrecht Institute, University Medical Center Utrecht and Princess Maxima Center, Utrecht, The Netherlands. [3]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA, USA. [4]Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA. [5]Moores Cancer Center, University of California, San Diego, La Jolla, CA, USA. [6]Department of Paediatrics, University of Cambridge, Cambridge, UK. [7]Academic Urology Group, Department of Surgery, Addenbrooke's Hospitals NHS Foundation Trust, University of Cambridge, Cambridge, UK. [8]Cell Screening Core, Department of Cell Biology, Center for Molecular Medicine, University Medical Centre Utrecht, Utrecht, The Netherlands. [9]Department of Surgery, Diakonessenhuis, Utrecht, The Netherlands. [10]Department of Pathology, Diakonessenhuis, Utrecht, The Netherlands. [11]Present address: Department of Gastroenterology, Keio University School of Medicine, Tokyo, Japan. [12]Present address: Core Life Analytics, Utrecht, The Netherlands. [13]These authors contributed equally: Sophie F. Roerink, Nobuo Sasaki, Henry Lee-Six. *e-mail: mrs@sanger.ac.uk; h.clevers@hubrecht.eu
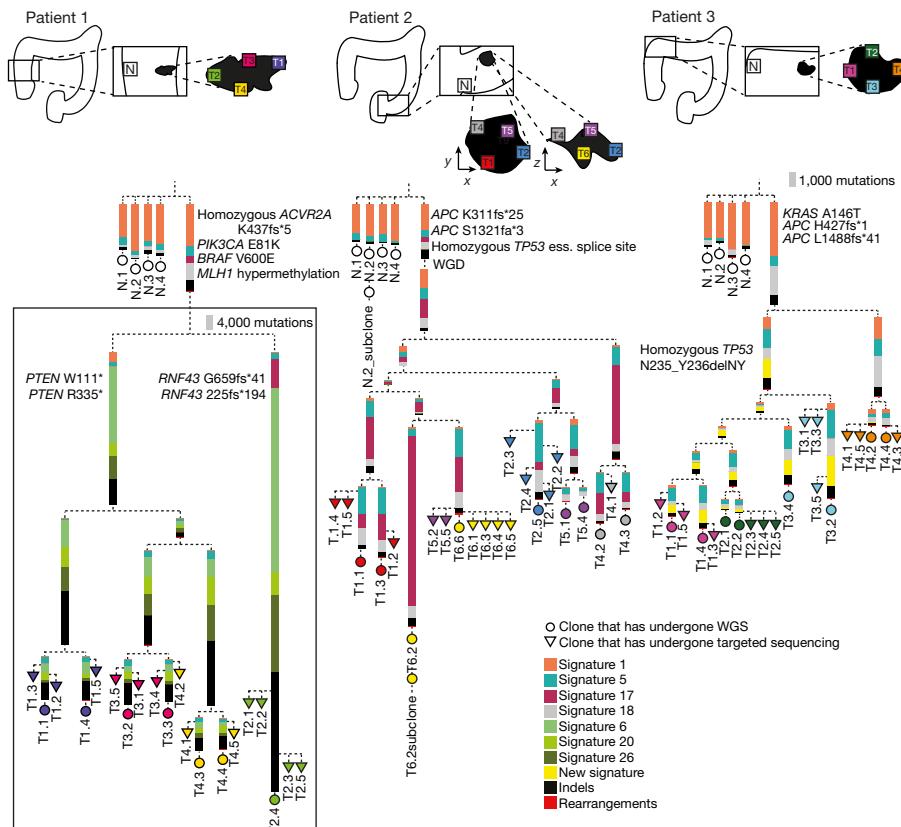
Fig. 1 | Mutation patterns during colorectal cancer evolution.
Multiregion sampling of each colorectal cancer is illustrated by coloured labels (T) and normal tissue sampling by white labels (N). Phylogenetic trees from three individuals have been constructed using somatic mutations in clonal colorectal organoids derived from normal and cancer cells. Organoids underwent whole-genome sequencing (circles) or targeted cancer gene panel sequencing (triangles). The lengths of branches are proportional to mutation numbers and each mutation type and mutation signature is indicated by a different colour. Driver mutations and a whole genome duplication (WGD) are indicated in the phylogenetic tree.

were identified by comparison with the sequences of DNA extracted from pieces of normal colorectal tissue. The overwhelming majority of somatic mutations identified in this way are likely to have occurred in vivo and not during in vitro culture (Extended Data Fig. 2c). Clonal organoids were also subjected to analysis of DNA methylation state at 470,000 CpG sites, RNA sequencing (RNA-seq), and were assessed for response to several anticancer therapeutics.

Phylogenetic trees of somatic mutations

Using the catalogues of somatic mutations from clonal organoids, we derived cell phylogenetic trees for each individual (Fig. 1a–c and Extended Data Figs. 3–6). The structures of the trees generally recapitulated the geographic origins of the clonal organoids within each cancer, with more closely related branches originating from the same tumour pieces. However, organoids within each piece continued to exhibit extensive genetic diversification: for example, in organoid clones isolated from tumour section P3.T3, at least 40% of mutations were not shared with other clones from this piece (Extended Data Fig. 1o).

In the trunk of the cancer cell phylogenetic tree of P1, we identified likely driver mutations in *BRAF* (V600E), *PIK3CA* (E81K) and *ACVR2A* (protein-truncating small indel). All cancer clones from this individual also showed microsatellite instability characteristic of DNA mismatch repair deficiency and hypermethylation of the *MLH1* promoter (the likely cause of this instability) (Extended Data Fig. 7). In addition, there were likely driver truncating mutations in *PTEN* and *RNF43* that were restricted to subsets of branches of the tree. In P2, two protein-truncating *APC* mutations and a homozygous splice site *TP53* mutation were present in the cancer trunk. In P3, a *KRAS* mutation (A146T) and two truncating *APC* mutations were present in the cancer trunk, and a *TP53* in-frame deletion was present in a subset of the branches. No driver mutations or *MLH1* methylation were observed in clonal organoids derived from normal colon epithelium from the three patients.

Mutation load in normal and cancer cells

A mean of 3,792 base substitutions was found in normal organoid clones derived from P1, 3,172 from P2 and 3,621 from P3 (Fig. 2), as previously reported[24]. The mean number of base substitutions in cancer-derived clones was higher in all three individuals: 72,398 in P1, 22,291 in P2 and 14,209 in P3. There were also substantial differences in the number of small indels and genome rearrangements. A mean of 227 small indels was observed in clones derived from normal colorectal epithelium from P1, 130 from P2 and 167 from P3. By comparison, the mean number of indels was 27,893 in cancer clones from P1, 1,485 from P2 and 2,021 from P3. There was a mean of one genome rearrangement in clonal organoids derived from normal colorectal epithelial cells contrasting with means in cancer-derived clonal organoids of 71 rearrangements from P1, 176 from P2 and 67 from P3. As the normal and cancer clones are derived from cells obtained from each individual at the same times, increases in somatic base substitution, small indel and genome rearrangement mutation rates are likely to have occurred in the lineages from fertilized egg to cancer cell, including in the two cancers that were proficient in DNA mismatch repair. Most of the additional mutation loads in cancer cells were acquired in the branches of the cancer phylogenetic tree rather than the trunk and therefore occurred following the last dominant clonal expansion within the cancer cell population.

Mutational signatures

We extracted mutational signatures and estimated the contributions of each signature to each segment of the phylogenetic trees. Eight base substitution mutational signatures were found (referred to according to the nomenclature in COSMIC http://cancer.sanger.ac.uk/cosmic/signatures).
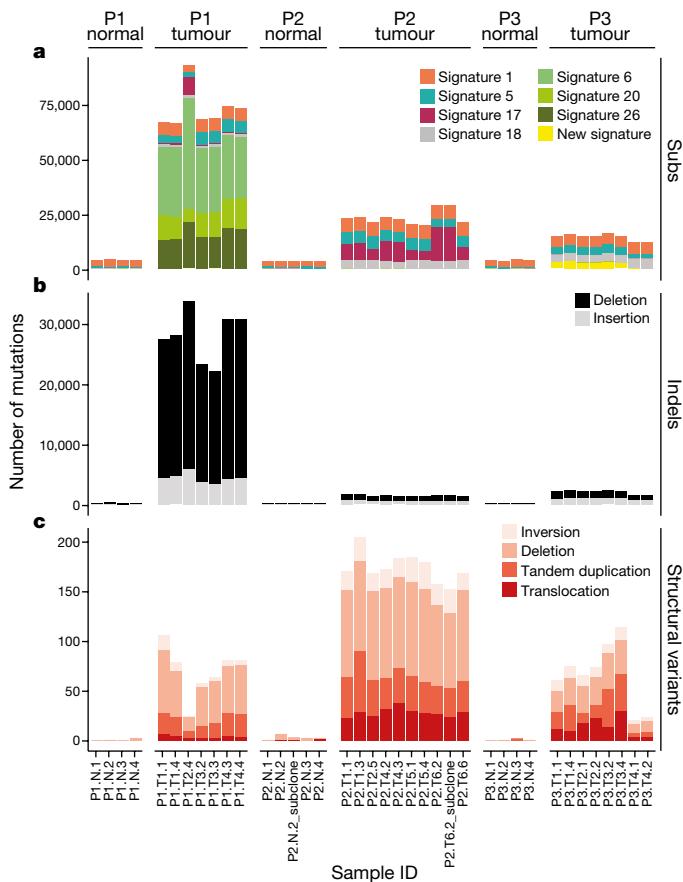
Fig. 2 | Total mutation burden in normal colorectal and colorectal cancer cells from three individuals. a, Substitutions have been further subdivided by contributions of mutational signatures. **b**, Indels have been subclassified as short insertions and deletions. **c**, Structural variations have been subdivided into deletions, inversions, tandem duplications and translocations.

These include the previously described signatures 1, 5, 6, 17, 18, 20 and 26 and a signature that has not been previously encountered[32] (see also Supplementary Notes section 5). Each mutational signature can be regarded as the outcome of a mutational process, which includes components of DNA damage or modification, DNA repair (or absence of it) and DNA replication, with each component potentially influencing the profile of the signature. Signature 1 is likely to result from deamination of 5-methylcytosine to thymine and has been reported to act in a 'clock-like' manner, with mutations accumulated continuously over the lifetime of an individual at different rates in different tissues. The number of signature 1 mutations is proposed to correlate with the number of mitotic divisions[33]. Signature 5 is of uncertain origin and also shows accumulation of mutations in a clock-like manner, with different rates in different tissues, although the rates do not correlate with those of signature 1[33]. Signatures 1 and 5 are found in most human cancers and probably in most normal cells[24,33].

Signature 1 dominated and, with signature 5, accounted for the large majority of mutations in normal colorectal stem cells[24] (Fig. 1). Signature 1 also dominated in the trunks of the cancer phylogenetic trees, presumably reflecting, at least in part, the long segment of normal cell lineage from the fertilized egg to the cell in which the first cancer driver mutation was acquired. However, in each of the three cancer trunks, signatures 17 and/or 18 also showed contributions (which were not detectable in the normal clonal organoids). Signatures 17 and 18 have been found in many cancer types and have uncertain mechanisms, although signature 18 may be related to DNA damage induced by reactive oxygen species and/or by deficiency of base excision repair.
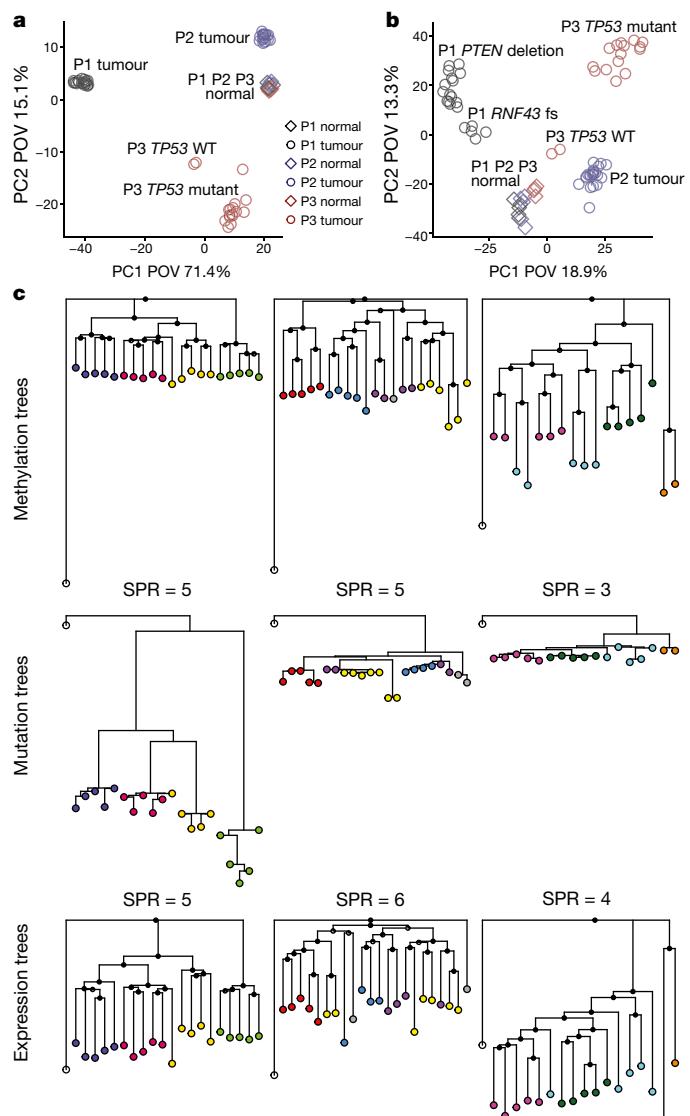

Fig. 3 | Diversification of methylation and transcriptome state during expansion of the neoplastic cell population. a, Clustering analysis of methylation state in each clonal organoid using 450 K Infinum arrays (n = 70 clones). PC, principal component; POV, percentage of variance. **b**, Clustering analysis of transcriptome state in each clonal organoid using RNA-seq (n = 73 clones). **c**. Phylogenetic trees based on methylation data (top), mutation data (middle) and expression data (bottom). Distances between mutation tree and methylation or expression tree topologies are expressed as subtree prune and regraft distance (SPR).

A different pattern of signature contributions was seen in the branches of the three cancer phylogenetic trees (Fig. 1). In P1, the mutations were predominantly of signatures 6, 20, 26 and indels; in P2, signatures 5, 17, 18, and indels; and in P3, signatures 5, 18, indels and a new signature predominantly characterized by T > G, T > A and T > C mutations at NTA and NTT trinucleotides (the mutated base is underlined; Extended Data Fig. 2d). The last signature occurred in all clones carrying a small in-frame deletion in *TP53*, but its relationship to this putative driver mutation is unknown. With regard to structural changes, P1 cancer clones carried deletions, inversions and tandem duplications, but few translocations, and their copy number profiles were relatively flat (Fig. 1 and Extended Data Figs. 5, 6). P2 cancer clones carried all types of rearrangement accompanied by changes in copy number profiles. In P3, *TP53* mutant clones carried abundant rearrangements of all types resulting in aberrant copy number states, whereas rearrangements were approximately tenfold less common in *TP53* wild-type clones with few copy number changes. The numbers of
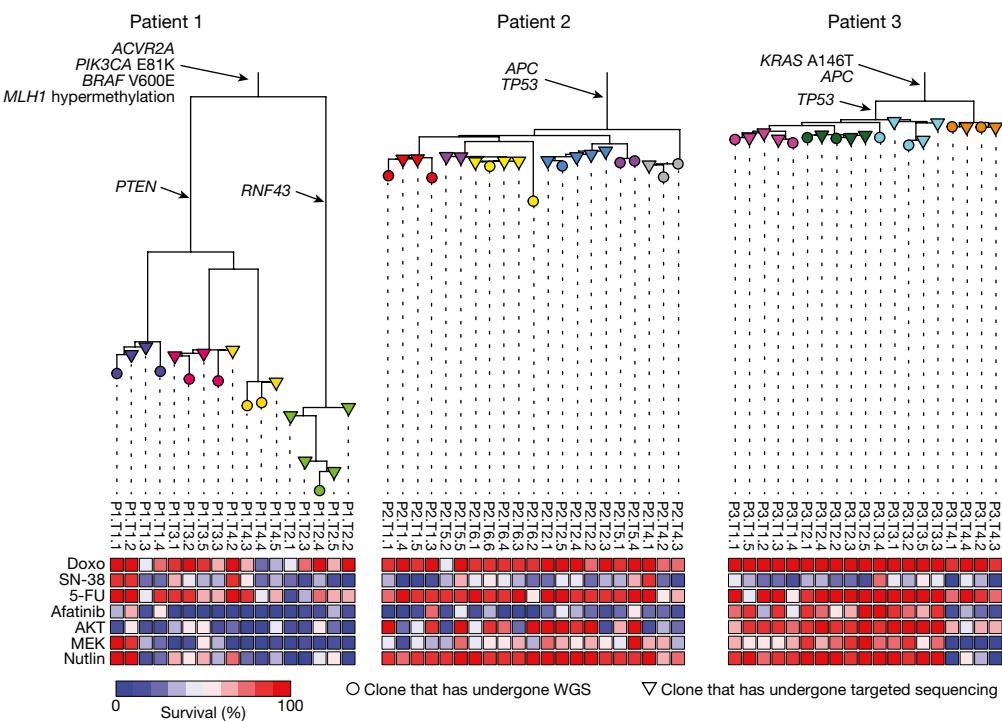
Fig. 4 | Individual clones show diverse responses to drugs commonly used in colorectal cancer treatment. Phylogenetic trees show the genetic structure of each tumour, with branch lengths representing mutation numbers. Coloured labels correspond to individual tumour segments as in Fig. 1. Mean survival in two independent experiments is displayed for exposure to doxorubicin (doxo), SN-38 (active metabolite of irinotecan), 5-FU, afatinib (an EGFR inhibitor), AKT inhibitor VIII (AKT), MEK1/2 inhibitor III (MEK) and nutlin-3a (nutlin; a stabilizer of TP53). Concentrations displayed represent intra- and inter-individual variation in the response to each drug; full dose response ranges are shown in Extended Data Fig. 10a.

mutations of several signatures differed markedly between individual branches, indicating varying contributions of mutational processes in different parts of the cancer.

There were more signature 1 mutations in each cancer organoid than in normal organoids from the same individuals (Fig. 1 and Extended Data Fig. 2e). Assuming that the clock-like correlation between the number of signature 1 mutations and number of mitoses undergone in normal cells is maintained at the same rate during neoplastic cell proliferation, we estimate that cancer cells from individual P1 have undergone 1.9 ($\pm$0.5) (s.d.) times as many mitoses as normal cells, cancer cells from individual P2 2.5 ($\pm$0.2) times as many, and from individual P3 1.7 ($\pm$0.2) times as many. An alternative explanation for the increase in signature 1 mutations in cancer cells is increased DNA methylation in cancer cells. However, cancer organoids were generally hypomethylated compared to normal cells (Extended Data Fig. 8b). In the distal branches of the phylogenetic trees (that is, during the most recent phases of cancer growth) the base substitution mutation rates per mitosis (as estimated by the total number of mutations divided by the number of signature 1 mutations) were markedly increased compared to normal cells (estimated 100-fold in P1, which is DNA mismatch-repair deficient, and tenfold in P2 and P3, which are mismatch-repair proficient). Thus, assuming that these estimates of past mitoses undergone are correct, the increases in base substitution, indel and genome rearrangement mutation rates over time also represent increases in mutation rates per mitosis.

Methylome and transcriptome

Epigenetic changes may be part of, and contribute to, the biological diversification of intra-tumour cell populations. To explore this possibility, we examined the methylation status of 470,000 CpG sites in the normal and tumour-derived clonal organoids. Organoids derived from normal stem cells from P1, P2 and P3 clustered together in principal component analyses, albeit with normal clones from each individual closer to each other than to normal clones from other individuals (Fig. 3

and Extended Data Fig. 8a). Clonal organoids from each colorectal cancer clustered together, with the exception of the two TP53 wild-type clones of P3, but separately from those derived from the other cancers and from normal organoids. Thus, the methylation states of normal colorectal stem cells from different individuals were relatively similar, but tumours from different individuals had developed divergent epigenetic states. For the P1 tumour, this conformed to the pattern of global hypermethylation previously termed CpG island methylator phenotype (CIMP)[25]. To investigate intra-tumour diversification of transcriptome state, we performed RNA-seq of normal colorectal epithelium and cancer-derived organoid clones. Clustering by gene expression profiles correlated well with clustering by methylation, with normal organoids from all individuals clustering together, while separate clusters existed for each cancer (Fig. 3b and Extended Data Fig. 9a). For each cancer we constructed phylogenetic trees based on methylation and gene expression (Fig. 3c). The topologies of the methylation and expression trees were remarkably similar to the mutation-based trees. Thus, diversification of methylation and transcriptome state occurred within each cancer and this was apparently heritable, stable and independent of the tumour microenvironment, as it persisted after organoid culture in vitro.

Diversification of drug responses

The clonal cancer organoids were exposed in vitro to a set of drugs used to treat colorectal cancer, including the chemotherapeutic agents 5-fluorouracil (5-FU), doxorubicin and 7-ethyl-10-hydroxycamptothecin (SN-38, the active metabolite of irinotecan), and the targeted agents afatinib (an epidermal growth factor receptor (EGFR) inhibitor), nutlin-3a (a stabilizer of TP53), a MEK1/2 inhibitor and an AKT inhibitor. Different organoids from the same cancer displayed substantial and reproducible differences in IC$_{50}$ values of up to 1,000-fold (Fig. 4 and Extended Data Fig. 10a, b), for both chemotherapeutic agents and targeted therapies. Some differences were attributable to particular somatic mutations. Notably, nutlin-3a exerted much greater growth

8888888888888888888

inhibition of *TP53* wild-type than mutant clones in P3 tumour organoids. Additionally, truncating mutations in *RNF43*, a recessive cancer gene encoding a negative regulator of the WNT pathway[34], rendered cells highly sensitive to the WNT secretion/porcupine inhibitor IWP2 (Extended Data Fig. 10c). The remaining variation in drug response did not, however, clearly relate to the geographical zones of origin or to the phylogenetic trees of each cancer. There were several examples of marked differences in drug sensitivity between closely related clones. For example, P2.T4.1 showed marked resistance to SN-38 compared to the other P2.T4 clones, whereas P3.T1.5 showed distinct sensitivity to 5-FU compared to all other clones from this individual (Fig. 4 and Extended Data Fig. 10a). The mechanism underlying this diversification in biological behaviour is unclear, but there was no obvious correlation with the degree of mutational diversification.

Discussion

Previous studies have addressed particular aspects of intra-cancer diversification by profiling the transcriptome, DNA copy number state and functional responses of individual cells[14–16,35]. To our knowledge, this is the first systematic and integrated analysis at genetic, epigenetic, transcriptomic and functional levels of multiple single-cell-derived clones from human cancers to incorporate high-quality and comprehensive description of essentially all somatic mutations present in multiple single cells. All three cancers studied, including the two DNA mismatch-repair proficient cases, clearly exhibited higher mutation burdens than normal colorectal stem cells. These are likely to result from increased mutation rates experienced during the lineages from fertilized egg to colorectal cancer cell. More comparisons of normal and colorectal cancer cells, and similar comparisons for other classes of cancer, are required for corroboration but it seems likely that increases in somatic mutation rates are common during the development of human cancers. These increases are predominantly due to recruitment of mutational processes that are inactive or marginally active in normal cells, and which dominate at later stages in the evolution of the cancer cell population. The roles of these processes in generating driver mutations, however, are unclear, as they may have started before or after acquisition of the early driver mutations in the trunks of the cancer phylogenetic trees[5]. The mechanisms underlying the increases in mutation rate in the DNA mismatch-repair proficient cancers are, for the most part, unknown. These increases may be due to somatic genetic or epigenetic changes (although these are not currently obvious), to metabolic stress attendant upon the elevated mitotic rate and other features of the neoplastic state, or to effects of the cancer cell microenvironment. Alongside intra-tumour mutational diversification, diversification of methylation state, transcriptome state and biological responses to therapeutics occur. While some methylation and transcriptome changes that occurred in vivo may have been lost during in vitro growth we were able to capture methylation and transcriptome changes that followed the evolution of the cancer through the mutational phylogenetic tree, which appeared to be stable and, at least partly, independent of the tumour cell microenvironment, as they persisted after cells were removed from the tumour. Diversification of methylation and transcriptome states and of drug responses are likely to result partly from driver mutations in cancer genes, but other, currently unknown, genetic and/or epigenetic mechanisms may be involved[36]. Future studies analysing the genomes, methylomes and transcriptomes of primary cells of a cancer will be needed to reveal all genetic, epigenetic and transcriptional variation occurring between cancer cells in vivo. Nevertheless, this study has shown the strength of the organoid system in stably retaining these characteristics and enabling functional assays on clones derived from individual cells. The analysis indicates that all three colorectal cancers contained cells resistant to most of the drugs commonly used to treat the disease. Differential drug responses between clones that are closely related both genetically and epigenetically suggest that resistance mechanisms can arise late in tumorigenesis.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0024-3.

1. Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
2. de Bruin, E. C. et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* **346**, 251–256 (2014).
3. Gundem, G. et al. The evolutionary history of lethal metastatic prostate cancer. *Nature* **520**, 353–357 (2015).
4. Sottoriva, A. et al. A Big Bang model of human colorectal tumor growth. *Nat. Genet.* **47**, 209–216 (2015).
5. Uchi, R. et al. Integrated multiregional analysis proposing a new model of colorectal cancer evolution. *PLoS Genet.* **12**, e1005778 (2016).
6. Suzuki, Y. et al. Multiregion ultra-deep sequencing reveals early intermixing and variable levels of intratumoral heterogeneity in colorectal cancer. *Mol. Oncol.* **11**, 124–139 (2017).
7. Navin, N. E. The first five years of single-cell cancer genomics and beyond. *Genome Res.* **25**, 1499–1507 (2015).
8. Gawad, C., Koh, W. & Quake, S. R. Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* **17**, 175–188 (2016).
9. Leung, M. L. et al. Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res.* **27**, 1287–1299 (2017).
10. Brocks, D. et al. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Reports* **8**, 798–806 (2014).
11. Oakes, C. C. et al. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov.* **4**, 348–361 (2014).
12. Mazor, T. et al. DNA methylation and somatic mutations converge on the cell cycle and define similar evolutionary histories in brain tumors. *Cancer Cell* **28**, 307–317 (2015).
13. Caiado, F., Silva-Santos, B. & Norell, H. Intra-tumour heterogeneity — going beyond genetics. *FEBS J.* **283**, 2245–2258 (2016).
14. Stevens, M. M. et al. Drug sensitivity of single cancer cells is predicted by changes in mass accumulation rate. *Nat. Biotechnol.* **34**, 1161–1167 (2016).
15. Dubach, J. M. et al. Quantitating drug–target engagement in single cells *in vitro* and *in vivo*. *Nat. Chem. Biol.* **13**, 168–173 (2017).
16. Tirosh, I. et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309–313 (2016).
17. Macaulay, I. C. et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **12**, 519–522 (2015).
18. Li, H. et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* **49**, 708–718 (2017).
19. Angermueller, C. et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229–232 (2016).
20. Meyer, M. et al. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proc. Natl Acad. Sci. USA* **112**, 851–856 (2015).
21. van de Wetering, M. et al. Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933–945 (2015).
22. Sato, T. et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772 (2011).
23. Behjati, S. et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425 (2014).
24. Blokzijl, F. et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016).
25. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
26. Tomlinson, I. P., Novelli, M. R. & Bodmer, W. F. The mutation rate and cancer. *Proc. Natl Acad. Sci. USA* **93**, 14800–14803 (1996).
27. Loeb, L. A. Human cancers express a mutator phenotype: hypothesis, origin, and consequences. *Cancer Res.* **76**, 2057–2059 (2016).
28. Tomasetti, C. & Vogelstein, B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**, 78–81 (2015).
29. Navin, N. E. Cancer genomics: one cell at a time. *Genome Biol.* **15**, 452 (2014).
30. Sieber, O. M., Heinimann, K. & Tomlinson, I. P. M. Genomic instability—the engine of tumorigenesis? *Nat. Rev. Cancer* **3**, 701–708 (2003).
31. Snippert, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
32. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
33. Alexandrov, L. B. et al. Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015).
34. Koo, B.-K. et al. Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**, 665–669 (2012).
35. Dalerba, P. et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127 (2011).
36. Landau, D. A. et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* **26**, 813–825 (2014).

METHODS

Human tissues. Tissue material was obtained from The Diakonessen Hospital, Utrecht. From the resected colon segment, both normal and tumour tissues were isolated. The isolated tumour tissue was subdivided into 4–5 segments. Normal tissue was taken at least 5 cm away from the tumour. All samples were obtained with informed consent and the study was approved by the ethical committee of The Diakonessen Hospital, Utrecht.

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Human organoid culture. Human normal and tumour colon organoids were established and maintained as described from isolated colonic epithelium[21,22]. In brief, long-term normal colonic organoid culture required human intestinal stem cell medium (HISM) composed of advanced DMEM/F12 (AdMEM) with penicillin/streptomycin, 10 mM HEPES, 1×GlutaMAX, 1×B27 (Invitrogen) and 1 μM N-acetylcysteine (SIGMA), supplemented with 50 ng ml^{-1} human recombinant EGF (Peprotech), 0.5 μM A83-01 (Tocris), 3 μM SB202190 (SIGMA), 1 μM nicotinamide (SIGMA), 10 nM prostaglandin E2,Wnt3A-conditioned medium (CM) (50% final concentration), Noggin-CM (10% final concentration), and R-Spondin1-CM (10% final concentration). Tumour organoids were cultured in medium containing only EGF, Noggin-CM, R-Spondin1-CM and A83-01.

Establishment of clonal organoids. For clonal organoids from normal crypts, isolated single crypts were embedded in 10 μl Matrigel and cultured in HISM medium. For clonal tumour organoid cultures, tumour cell suspensions were cultured for 7–14 days in HISM without Wnt3A-CM. Then, 10–15 individual organoids were picked and separately dissociated into single cells by TryPLE express (Thermo Fisher), washed and suspended in AdMEM containing propidium iodide (PI). Forty-eight single cells were sorted into tumour organoid medium (HISM plus 10 μM ROCK inhibitor Y-27632 (Tocris BioScience); no Wnt-CM) from each tumour organoid. Sorting was based on FCS area/FCS peak and PIneg/FCS area using a Moflo machine (Beckman Coulter). Sorted cells were spun down at 1,000g at 4 °C for 5 min, after which single cells were each embedded into 10 μl of basement membrane extract (BME, Amsbio) and seeded into 96-well plates at a ratio of 1 cell per well. The gel was left to solidify in a 37 °C incubator after HISM (no Wnt3A-CM) was added. Y-27632 was added to the medium for the first week after sorting. For each original tumour organoid, a single clonal organoid was selected and expanded for further study and for preparing frozen stocks. Culturing times and plating efficiencies are listed for each organoid in Supplementary Datafile S1.

Histology procedures. Tissues were fixed in 4% formaldehyde solution overnight and embedded in paraffin. Sections were subjected to haematoxylin and eosin (H&E) and immunohistochemistry staining. The Ki67 antibody (MONX10283, Monosan) was used at 1:250 dilution.

Organoid CellTiter-Glo viability assay. Tumour organoids were cultured for 5–10 days after being trypsinized into single cells in HISM without Wnt-3a-CM. The organoids were mechanically dissociated by pipetting before being resuspended in 5% BME/growth medium (25 × 10^3 organoids per ml). Before seeding, 10 μl BME was dispensed into 384-well plates, and then 30 μl growth medium containing organoids was dispensed into each plate (at about 750 organoids per well). Drug screening was carried out using nutlin-3a, afatinib, MEK1/2 inhibitor III, AKT inhibitor VIII, 5-FU, doxorubicin, and SN-38. Drug dilutions were performed in two series: 1) stepwise 2 fold-dilutions from 20 μM to 19.5 nM; and 2) stepwise 2 fold-dilutions from 15 μM to 29.3 nM. The measurements for these two dilution series were combined into a single curve. All drugs were dispensed by a HP-D300 automated liquid dispenser (TECAN). Samples were incubated for 6 days at 37 °C, and cell viability was measured by CellTiter-Glo 3D kit (Promega) on a SpectraMax M5e (Molecular Devices). Cell viability measurements were performed in duplicate wells for each clone. Survival ratios in drug-treated organoids were normalized to the average survival in a DMSO control. Each experiment was repeated on a different day. To assess variability between technical and biological replicates we calculated the area under the curve (AUC) for each survival curve. AUC values are calculated using the trapezoid method and are divided by the area covered by 100% survival on the y-axis and the range of the log$_{10}$ concentrations on the x-axis (Extended Data Fig. 10b).

DNA and RNA extraction. DNA and RNA were concomitantly extracted from frozen tissue samples or organoid cultures using AllPrep DNA/RNA minikit (Qiagen 80204).

Whole-genome sequencing. From each individual, 7–10 tumour-derived clones were selected for whole-genome sequencing (WGS), as well as 4–5 normal-derived clones. We generated paired end sequencing reads (150 bp) using Illumina XTEN machines, resulting in about 30× coverage per sample. Sequences were aligned to the human reference genome (NCBI build37) using BWA-MEM. Sequencing statistics are listed in Supplementary Datafile S2.

Cancer gene panel sequencing. All WGS sequenced clones and 40 additional tumour-derived clones were subjected to targeted sequencing. An in-house developed cancer gene panel (CGPv3) was used, designed to pull down 360 genes that are known or suspected to play a role in cancer[3,37]. The panel targets genes from the Cancer Gene Census (COSMIC), genes recurrently amplified or over-expressed in cancer and candidate cancer genes such as kinases from the MAP kinase signalling pathway. We performed custom RNA bait design following the manufacturer's guidelines (SureSelect, Agilent) and previously described workflows to create pulldown libraries from native genomic DNA[3,37]. Samples were multiplexed on flow cells and subjected to paired end sequencing (75-bp reads) using Illumina HiSeq2000 machines, resulting in more than 700× coverage for the target design for tumour-derived samples and more than 2,000× coverage for normal tissue-derived samples. Sequences were aligned to the human reference genome (NCBI build37) using BWA-align.

RNA sequencing. From each individual, RNA was isolated from all tumour and normal derived organoid clones and subjected to RNA-seq analysis. RNA-seq libraries were prepared according to previously described workflows and sequenced on Illumina HiSeq2000 machines[38]. Between four and seven barcoded samples were pooled per library. Sequenced reads were aligned to the human reference genome (NCBI build37) with Bowtie/TopHat.

Methylation arrays. Infinium HumanMethylation450 BeadChip arrays were used to characterize the methylation status of more than 450,000 CpG sites for all clones.

Mutation discovery. All somatic changes in whole genome and targeted data were analysed with mutation calling pipelines developed in house (available at https://github.com/cancerit).

Substitutions. Single-base somatic substitutions were identified using CaVEMan[39] and a number of post-processing filters were applied. For each patient, the only germline reference available was healthy colorectal tissue more than 5 cm from the tumour, consisting of epithelial and connective tissue. In order to allow the discovery of a field effect[40,41] that might have spread into the matched normal sample, we ran CaVEMan using an unmatched normal reference. Germline SNPs were removed by comparison to a panel of 75 unrelated normal samples. Additional post-processing filters were applied to these mutation calls as described[42]. For XTEN/BWA-mem aligned data we added two filters to the pipeline for the median alignment score (ASMD) ≥ 140 and the clipping index (CLPM) = 0, meaning that fewer than half of the reads should be clipped.

We then tabulated the number of mutant and wild-type reads for every mutation discovered in every sample, including the adjacent colorectal tissue. We considered only mutations that were covered by ten reads in all related clones, and mutations that were seen on at least two reads in each direction. As the adjacent normal colorectal tissue is not entirely composed of epithelium we reasoned that if there were a field effect the somatic mutations in this tissue should not be fully clonal. We therefore deducted germline mutations on the basis that they were fully clonal in the bulk normal, while mutations that were subclonal in the bulk were not removed from the analysis. To define a mutation as subclonal in the bulk, the probability of finding the observed number of mutant reads or fewer given the sequencing coverage had to be less than 0.005, based on the binomial distribution with a probability of 0.5 for autosomes. Mutations that failed to meet this criterion were considered to be germline and were removed.

Indels. Indels were called using Pindel[43,44] using the adjacent bulk colorectal tissue as a matched normal. Post processing filters were the same as for substitutions except that ASMD ≥ 140 and CLPM = 0 were not used.

Copy number. Copy number profiles were constructed for WGS samples by ASCAT[45,46], using adjacent healthy colorectal tissue as a matched normal. Copy number profiles of WGS analysed samples were visualized with the plotHeatmap function of the R package 'copynumber'[47].

Rearrangements. Rearrangements were called using healthy adjacent colorectal tissue as a matched normal. Abnormally paired read pairs from WGS were grouped and filtered by read remapping using 'Brass' (https://github.com/cancerit/BRASS). Read pair clusters with 50% or more of the reads mapping to microbial sequences were removed. Candidate breakpoints were matched to copy number breakpoints defined by ASCAT within 10 kb. Rearrangements not associated with copy number breakpoints or with a copy number change of less than 0.3 were removed.

Phylogenetic tree construction. The phylogeny of single-cell-derived organoids for each patient was constructed from substitutions called in WGS data. For each patient, substitutions that had been discovered by CaVEMan in any organoid from that patient were called as present or absent in each organoid using the algorithm Shearwater[48]. This algorithm compares allele frequencies of variants to a background error model derived from sequencing thousands of samples from unrelated studies on the same platform. Sequencing errors are known to occur at different frequencies across different sites of the genome[48]. By obtaining a comprehensive view of the number of variant calls at each position in unrelated normal genomes, we built an error model for each nucleotide change for each position. This method has previously been used to find variants at a low frequency[49]. Here, we compared the observed frequency of each variant to our error model. After correcting for multiple testing, only variants that were significantly mutated over the error

model were kept, using a *q* value cutoff of 0.05. Although most true variants largely exceed this threshold, this procedure maximizes the chance of retaining variants in genomic regions that have undergone copy number changes, lowering the apparent variant allele fraction (VAF), since the Shearwater algorithm was designed to detect subclonal variants. As a further stringent filter to minimise false positive calls, variants had to be supported by at least three mutant reads to be considered by the algorithm. In this way every mutation called in an individual was genotyped as being present or absent in each sample. Phylogenies were constructed using this binary matrix of mutations present or absent in each sample. Private mutations were discarded from tree building as they are uninformative. A fake outgroup with no mutations called was generated for each individual. Phylogenies were constructed using the Phylip suite of tools[50]. The programme seqboot was used to generate 100 bootstrap replicates of each dataset by resampling the mutations with replacement. Phylogenies were then reconstructed for each bootstrap replicate by maximum parsimony using the Mix programme using the Wagner method, using the jumble = 10 option was used, randomizing the order of the input samples 10 times for each bootstrap replicate. Finally, the programme Consense was used to build a consensus of all the trees that had been built for each patient, using the majority rule (extended) option. This reports, for each node in the most parsimonious tree, how many of the trees that had been built contain a node that partitions the samples into the same two groups. All nodes in each tree that relate tumour samples to each other were supported by all bootstrap replicates (Supplementary Notes).

Assignment of somatic changes in WGS to the phylogenetic tree. *Substitutions.* Substitutions were called as present or absent in each organoid as described above. To assign these mutations to the tree, each branch of the tree was considered in turn. If a mutation was called in all the organoids that were descendants of a given branch, and in no organoids that were not descendants of the branch, mutations were assigned to that branch. Ignoring private mutations, which necessarily fit any tree, 97.7% of shared mutations fitted the tree structure from patient 1 perfectly, 89.7% fitted the tree from patient 2 perfectly, and 88.1% fitted the tree from patient 3 perfectly. The lower concordance with the tree for patients 2 and 3 reflects the increased copy number changes that have occurred in these phylogenies. Examination of the copy number state at loci where there were discordant mutations showed that the majority could be explained by deletions of those mutations in a subclone. Substitutions that did not fit the tree perfectly were therefore assigned to the most recent common ancestor of the samples in which they were called.

All substitution calls and their assignment to branches of the tree, as well as substitutions that did not fit the tree perfectly and their associated copy number states are listed in Supplementary Data file S4.

Indels. Indels were called as being present or absent in each sample based on a variant allele fraction (the proportion of mutant reads at a locus) cutoff. The variant allele fraction cutoff was chosen for each patient based on a histogram of the variant allele fraction to separate the sequencing noise distribution from the distribution of true mutations. Variant allele fraction cutoffs were chosen as 0.15 for patients 1 and 3, and 0.11 for patient 2. Indels were then assigned to branches of the tree that they fitted perfectly. Indels that were assigned to the tree, along with their assignments, as well as indels that did not fit the tree are provided in Supplementary Data file S4.

Rearrangements. The same rearrangement may be called in related samples with slightly different breakpoints. To identify rearrangements that had been sequenced in related clones as the same, both the upstream and the downstream breakpoints had to fall within 500 bp of each other. The majority of rearrangements fitted the tree. Visualization of discordant rearrangements using IGV[51] showed that often an overlapping rearrangement meant that the rearrangement was lost in a clone. All rearrangement calls that could and could not be assigned to the tree can be found in Supplementary Data file S4.

Timing substitutions and indels relative to a whole genome duplication (WGD). A whole genome duplication was observed in the trunk of the tumour for patient 2. *Substitutions.* For substitutions, we aimed first to obtain an accurate estimate of the timing of WGD in molecular time, and second to time as many substitutions as possible relative to the WGD in order to perform signature analysis on them.

To obtain an estimate of the timing of the WGD, we examined substitutions in regions with two copies of one allele and none of the other (as determined by ASCAT). In these regions, if a mutation occurred before the WGD on that allele, it will be at copy number 2. If it occurred afterwards, it will be at copy number 1. One hundred and eighty substitutions occurred at copy number 2, and 67 at copy number 1. As at least half of the mutations that occurred before the WGD have been lost in such regions (as there is loss of one allele), the WGD can be estimated to have occurred at 84% ((180 × 2)/(180 × 2 + 67)) of molecular time in the trunk of the tumour (95% confidence interval of 80.8–87.6% calculated by bootstrapping 10,000 times).

Second, we wanted to time substitutions in regions with a greater range of copy number states for mutational signature analysis. To do this, for every truncal substitution in every tumour clone from patient 2, the copy number segment (as called by ASCAT) in which that mutation fell was defined. Mutations could be timed only in samples in which there was a minor copy number of 0 and a major copy number greater than 1. Fortunately, because of the extensive copy number changes in this tumour, all mutations fell in a region that met these criteria in at least one sample. For a given mutation that fell in such a copy number segment in a given sample, the VAF in that sample of known germline single nucleotide polymorphisms (SNPs) that fell in that segment (that necessarily occurred before the WGD) and the VAF of somatic mutations assigned to branches further down the tree (that necessarily occurred after the WGD) was examined. If, in a given sample, a mutation had a VAF greater than 90% of the VAFs of the mutations that were known to occur further down tree it was considered to have occurred before the WGD, whereas if it had a VAF less than 90% of the VAFs of the SNPs it was considered to have occurred after the WGD. If there was any overlap between the 90th percentiles of the SNPs and the later mutations, or if the mutation fitted neither of these criteria, it was considered uninformative and was not used in the signature analysis. This accounted for 9,094 mutations (out of a total of 12,623 assigned to the trunk) that were not used in signature analysis. There is no reason to believe that mutations that were excluded for these reasons should be attributable to different mutational signatures than those that could be included, and indeed their trinucleotide mutation contexts are similar (data not shown). For each mutation, then, the number of samples in which it had been counted before and after the WGD was tallied. If a mutation was called as occurring before the WGD in some samples and after the WGD in others, the mutation was designated as conflicting and excluded from the analysis. Eighty-two mutations fell into this category, and the remaining 3,447 could be timed unambiguously relative to the WGD and used in the signature analysis. In Fig. 1 we extrapolated the preWGD and postWGD fractions and their relative signature components to all mutations identified in the clonal trunk of P2. Mutations that were included in the signature analysis, those that were excluded as being uninformative, and those that were excluded as being conflicting, are reported in Supplementary Datafile S4.

Indels. For indels, we simply aimed to estimate the proportion that occurred before and after the WGD, and so for this analysis we restricted ourselves to regions of the genome with copy number 2 + 0. An analogous approach to timing substitutions was taken, although rather than considering the distribution of germline indels and indels further down the tree, a hard VAF cutoff of 0.85 (which separated the bimodal distribution of indel VAFs in these loci) was used to define mutations as occurring before or after the WGD.

Rearrangements. We timed rearrangements relative to the WGD in patient 2 by using the copy number step associated with deletions and tandem duplications in the trunk of this tumour, as determined by inspection of the change in read counts at breakpoints. The ratio of tandem duplications and deletions that had occurred before rather than after the WGD was extrapolated to give a ratio for all the rearrangements in the trunk, assuming that the relative proportion of different rearrangement classes stayed the same after the WGD.

Driver mutations. Driver mutations in TP53 and APC were timed relative to the WGD in patient 2. The TP53 mutation was at VAF 1 in a region that was 2 + 0 in all samples, indicating that it occurred before the WGD. There were mutations in both alleles (which we will call mutation 1 and mutation 2) of *APC*. P2.T4.2 and P2.T5.1 both had the *APC* locus called as 2 + 2, and both mutations were at VAF 0.5. P2.T1.1, P2.T1.3, and P2.T6.2 were 2 + 1 in the *APC* region. Mutation 1 was at VAF 0.67 and mutation 2 at 0.33. In P2.T2.5 the region was also called as 2 + 1, but mutation 1 was at VAF 0.33 and mutation 2 at VAF 0.67. This shows biallelic inactivation of *APC* before the WGD.

Assignment of samples to the tree based on targeted sequencing. Samples for which both WGS and targeted sequencing were available were used to estimate the sensitivity in the targeted data for finding substitutions that were identified in each branch by the WGS data. The targeted capture was designed against 360 cancer genes; in addition, all off-target reads that covered substitutions identified by WGS were considered. For example, in clone P1.T1.1, a fraction of 0.09 of all substitutions in the ultimate branch P1.T1.1 was found in the targeted data. Samples for which only targeted sequencing was available were then assigned to the ultimate branch of the tree with which they shared most substitutions. For example, clone P1.T1.4 shared a fraction of 0.04 with branch P1.T1.1 and negligible fractions with other branches. To estimate the time point at which P1.T1.4 branched off, we divided the shared fraction 0.04 in clone P1.T1.4 by the sensitivity in clone P1.T1.1 0.09. Thus, we estimated that P1.T1.4 shares 0.04/0.09 = 43% of mutations with branch P1.T1.1. The proportion of mutations shared with each branch are listed in Supplementary Notes section 2.

Driver analysis. To classify driver events in substitutions, indels and rearrangements we used the following criteria: 1) deleterious mutations in genes identified in CRC by TCGA[25]; 2) all other known oncogenes carrying a canonical

activating mutation; and 3) tumour suppressor genes with loss of function, and/or carrying two deleterious mutations. A more inclusive approach for identifying functional mutations is listed in Supplementary Datafile S3 and described in the Supplementary Notes.

Mutational signature analysis. Signature extraction based on non-negative matrix factorization was performed as previously described[32,52]. Mutations in trinucleotide context were grouped according to branches of the phylogenetic tree or according to sample. Datasets were combined with data from 560 breast cancer genomes to increase performance of the NNMF procedure[53].

Expression analysis. Clustering analyses were based on FPKM values calculated with the Cufflinks algorithm[54]. To select informative genes for clustering we applied the following filters: FPKM > 1; coefficient of variation across all samples > 0.7 or absolute difference > 5. Subsequently, FPKM values were $\log_2$ transformed and normalized. Normalization across samples was applied by subtracting the median expression value from individual expression values. Normalization across genes was applied by subtracting the gene's median expression from individual expression values. Normalized values were subjected to principle component analysis (PCA).

For each tumour clone we calculated a set of genes differentially expressed compared to all normal clones pooled together. Genes were called as differentially expressed if they had an FDR-corrected P value less than 0.05, resulting from a likelihood ratio test using a negative binomial generalized linear model fit with the R package 'edgeR'[55]. Raw counts were input into the edgeR model, along with normalization offsets calculated using the TMM method[56]. To construct the expression-based phylogenetic trees, we calculated Euclidean distances based on all genes that were differentially expressed in at least one tumour clone from that individual. Trees were inferred by the minimum evolution method, with the fast-me.bal function in the R-package 'ape'[57].

Methylation analysis. Methylation arrays were processed using the R package minfi[58]. We excluded three samples that failed standard QC metrics (more than 1,000 failed probes). We then excluded any probe from the analysis if it contained a variant identified in one of the samples, had a detection P value $> 1 \times 10^{-10}$ in > 10% samples or one sample with $P > 0.01$, occurred at a location known to cross-hybridize with another genomic location, or where there was a known SNP at the CpG targeted by the probe. The remaining probes were then normalized either using the 'preprocessRaw' function when comparing all samples together or using the SWAN quantile-normalization method when comparing clones from a particular tumour[59]. The latter method is appropriate when the number of methylated probes is expected to be roughly constant across all samples, which was the case for each tumour. For all comparisons, M values ($\log_2$ ratio of methylated to unmethylated probes) were calculated from the normalized probe intensities. Samples were clustered using PCA. For computational reasons, we used only the 1,000 most variable probes for PCA.

Probes that were differentially methylated between tumour and normal cells were identified using an F-test with variance shrinking and a false discovery rate of 0.01 via the 'dmpFinder' function comparing each tumour clone to all normal derived clones[58]. To construct the methylation-based phylogenetic trees, we calculated Euclidean distances based on probes that were differentially methylated in at least one tumour clone from that individual. For computational reasons, we used a q-value cutoff of 1×10^{-8} for selection of informative probes. Trees were inferred by the minimum evolution method, with the fast-me.bal function in the R-package 'ape'[57].

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Code availability. Mutation calling pipelines developed in house are available at https://github.com/cancerit. The Shearwater algorithm for deriving a background error model is available at: https://www.bioconductor.org/packages/devel/bioc/vignettes/deepSNV/inst/doc/shearwaterML.html. The software for signature analysis used in this manuscript is available at: https://www.mathworks.com/matlabcentral/fileexchange/38724-wtsi-mutational-signature-framework.

Custom R scripts developed for the analyses and visualizations in this manuscript are available from the authors on request.

Data access. Sequencing data have been deposited at the European Genome-Phenome Archive (http://www.ebi.ac.uk/ega/) under accession numbers EGAS00001000869 (targeted sequencing), EGAS00001000985 (RNA-seq) and EGAS00001000881 (WGS). RNA sequencing data of these organoid clones has also been described elsewhere[60].

37. Yates, L. R. et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).
38. Iorio, F. et al. A landscape of pharmacogenomic interactions in cancer. *Cell* **166**, 740–754 (2016).
39. Jones, D. et al. cgpCaVEManWrapper: Simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. *Curr. Protoc. Bioinformatics* **56**, 15.10.1–15.10.18 (2016).
40. Lochhead, P. et al. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Mod. Pathol.* **28**, 14–29 (2015).
41. Luo, Y., Yu, M. & Grady, W. M. Field cancerization in the colon: a role for aberrant DNA methylation? *Gastroenterol. Rep. (Oxf.)* **2**, 16–20 (2014).
42. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
43. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).
44. Raine, K. M. et al. cgpPindel: identifying somatically acquired insertion and deletion events from paired end sequencing. *Curr. Protoc. Bioinformatics* **52**, 15.7.1–15.7.12 (2015).
45. Raine, K. M. et al. ascatNgs: identifying somatically acquired copy-number alterations from whole-genome sequencing data. *Curr. Protoc. Bioinformatics* **56**, 15.9.1–15.9.17 (2016).
46. Van Loo, P. et al. Allele-specific copy number analysis of tumors. *Proc. Natl Acad. Sci. USA* **107**, 16910–16915 (2010).
47. Nilsen, G. et al. Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics* **13**, 591 (2012).
48. Gerstung, M., Papaemmanuil, E. & Campbell, P. J. Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics* **30**, 1198–1204 (2014).
49. Martincorena, I. et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015).
50. Felsenstein, J. PHYLIP — Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166 (1989).
51. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
52. Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Reports* **3**, 246–259 (2013).
53. Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).
54. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols* **7**, 562–578 (2012).
55. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
56. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
57. Desper, R. & Gascuel, O. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *J. Comput. Biol.* **9**, 687–705 (2002).
58. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
59. Maksimovic, J., Gordon, L. & Oshlack, A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* **13**, R44 (2012).
60. Borten, M. A., Bajikar, S. S., Sasaki, N., Clevers, H. & Janes, K. A. Automated brightfield morphometry of 3D organoid populations by OrganoSeg. *Sci. Rep.* (in the press).
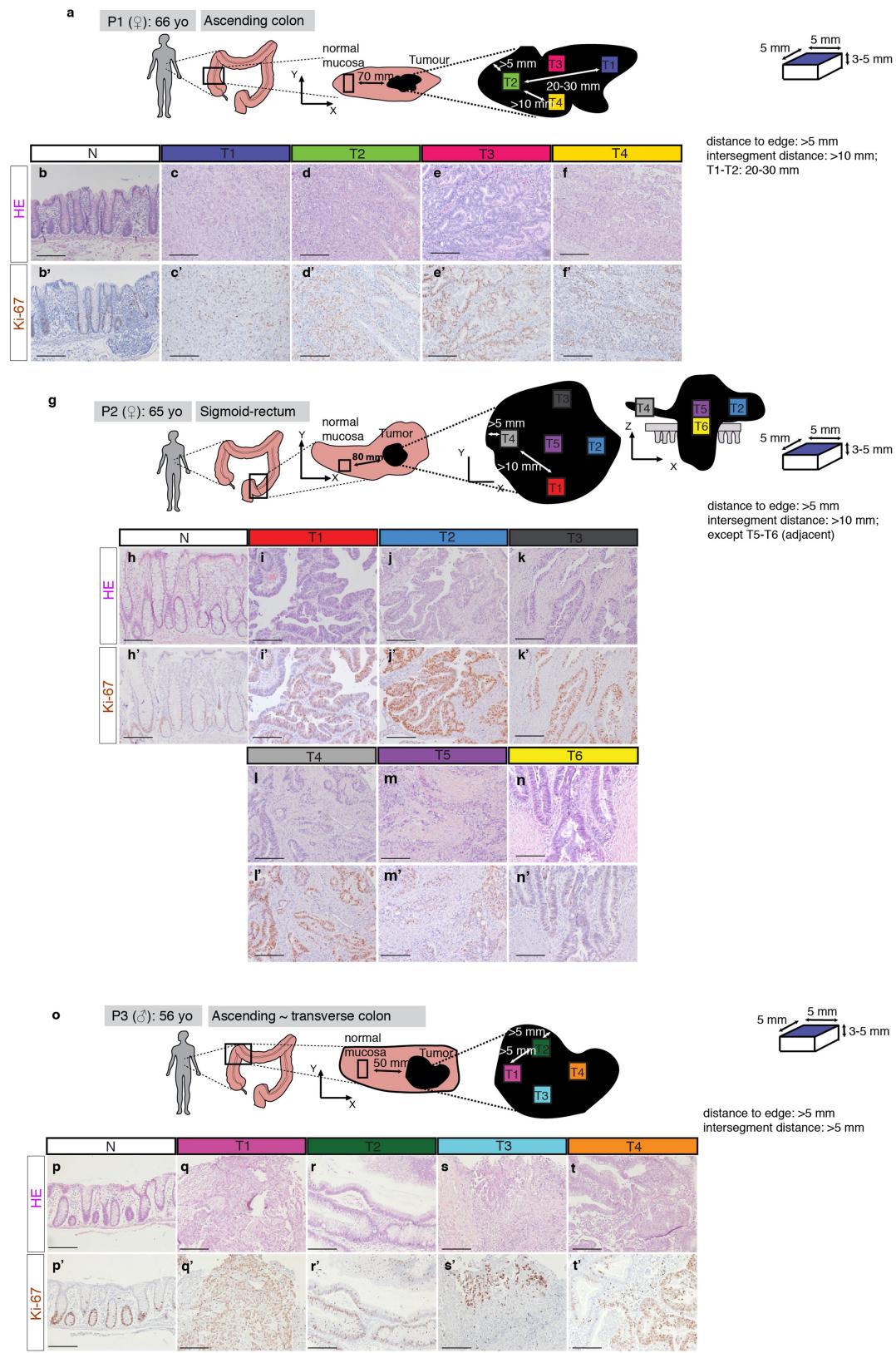
Extended Data Fig. 1 | Origin of clonal organoids analysed in this study. Specimens were derived from the ascending colon of a 66-year-old woman (**a–f**), sigmoid rectum of a 65-year-old woman (**g–n**) and ascending transverse colon of a 56-year-old man (**o–t**), respectively. From each tumour, 4–6 segments were resected (sized 5 × 5 × 3–5 mm. All sections except T3 from P2 resulted in viable clonal organoids. **b–f**, **h–n**, **p–t**, Haematoxylin and eosin staining and Ki67 immunohistochemistry show cell morphology for individual tumour sections. Scale bars: 200 μm.
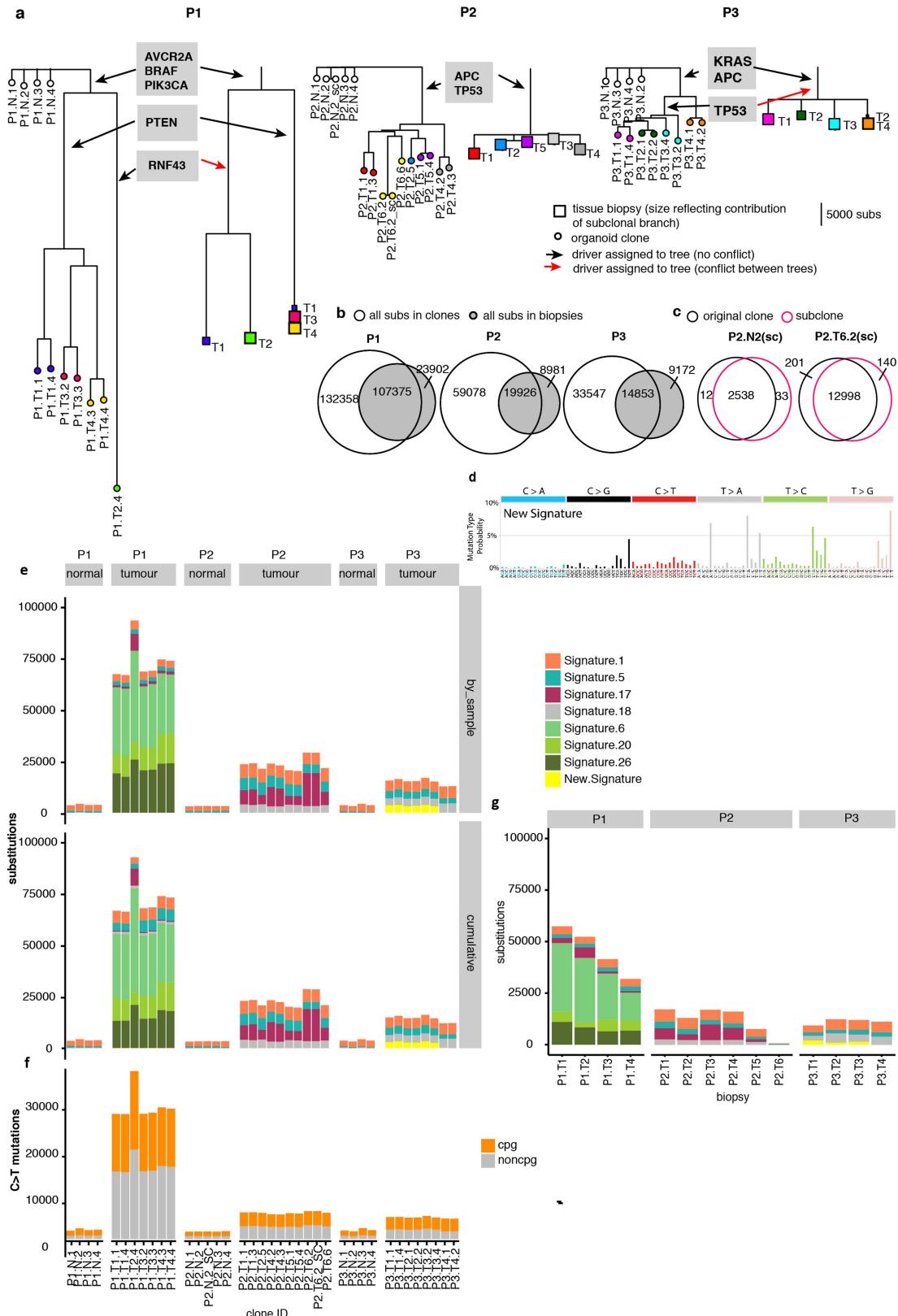
Extended Data Fig. 2 | Substitution analysis. a, Comparison of phylogeny reconstructions from WGS analysis of clonal organoids (left) and subclonal analysis of the original tissue biopsies (right) from individuals P1–P3. The analysis of clonal organoids allows a very detailed phylogeny, exact placement of driver mutations and analysis of cell-to-cell differences. **b**, Venn diagrams depicting overlap between substitutions identified by the organoid approach and the tissue biopsy approach. **c**, Venn diagrams depicting overlaps between clones P2.N3 and P2.T6.2 and their respective subclones (see Methods). Only a small proportion of the total mutations is added during culturing in both normal and tumour organoids. **d**, New

signature identified in this study in tumour organoid samples from P3, characterized by T > G, T > A and T > C mutations at N$\underline{T}$A and N$\underline{T}$T trinucleotides (mutated bases underlined). **e**, Contribution of each of the identified mutation signatures to individual samples. Top (by_sample), results of signature extraction from all substitutions identified in each sample (Supplementary Notes). Bottom, proportion in each sample derived by adding up proportions in the branches of the phylogenetic tree that make up that sample (identical to Fig. 1). **f**, Numbers of C > T mutations by CpG context. **g**, Signature analysis of substitutions identified in the original tissue biopsies.
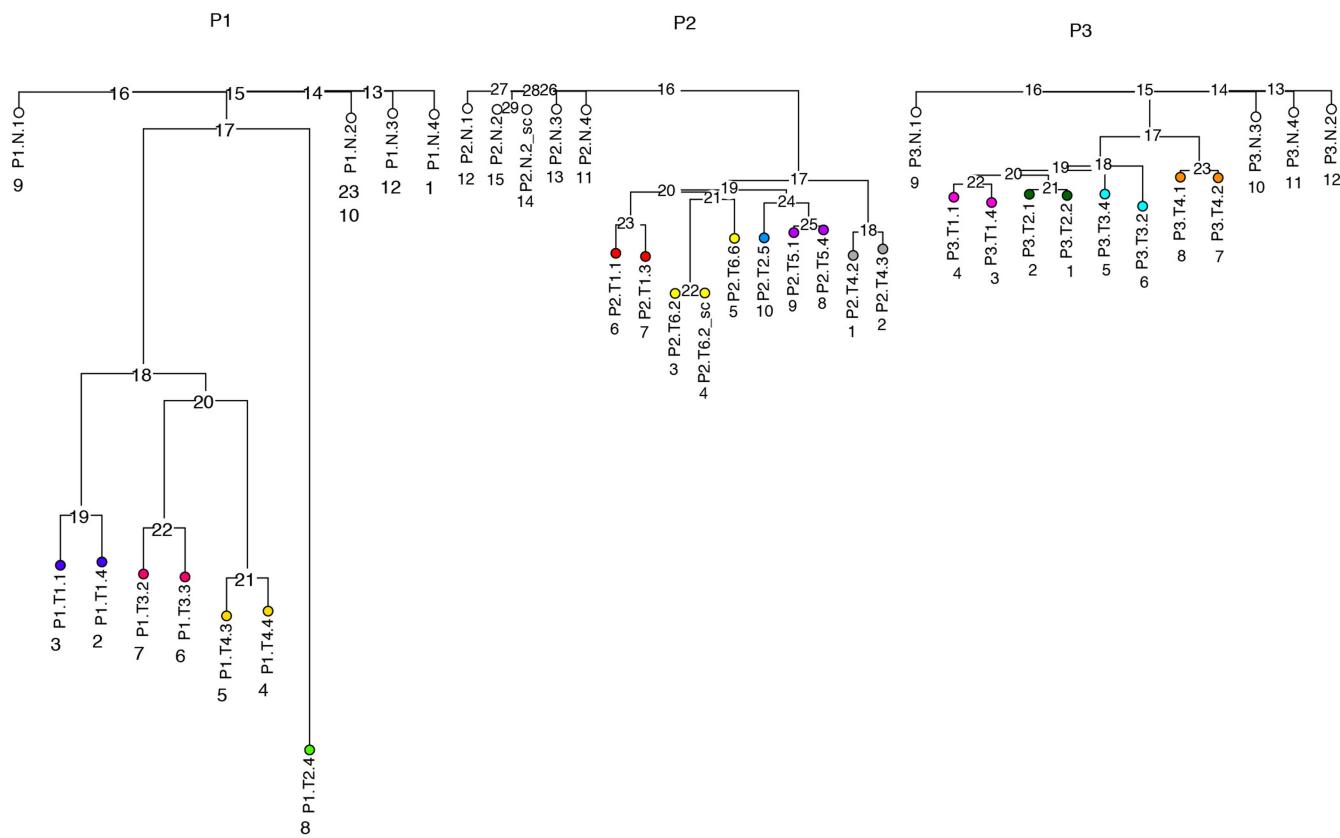
Extended Data Fig. 3 | Phylogenetic trees for clones that have been analysed by WGS. Branch lengths represent total mutation numbers; labels of nodes and tips in the tree correspond with labels in Supplementary Data files S3–S5.
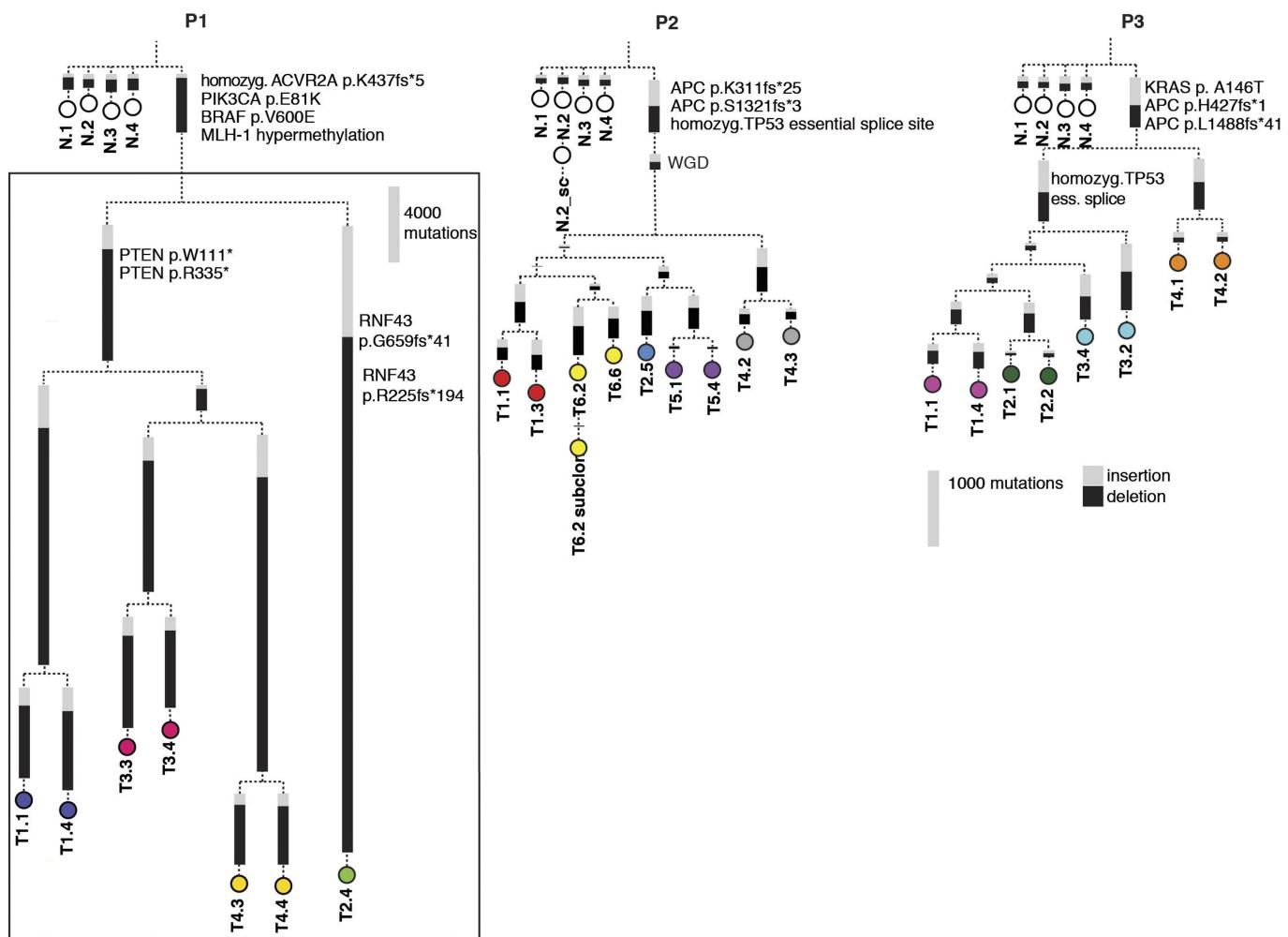
Extended Data Fig. 4 | Phylogenetic trees for indels. Phylogenies for three individuals with branch lengths representing indel numbers, further subdivided in insertions and deletions. Boxed area for P1 shows the high number of indels in this patient, who displays microsatellite instability (MSI) in all tumour clones in a different scale.
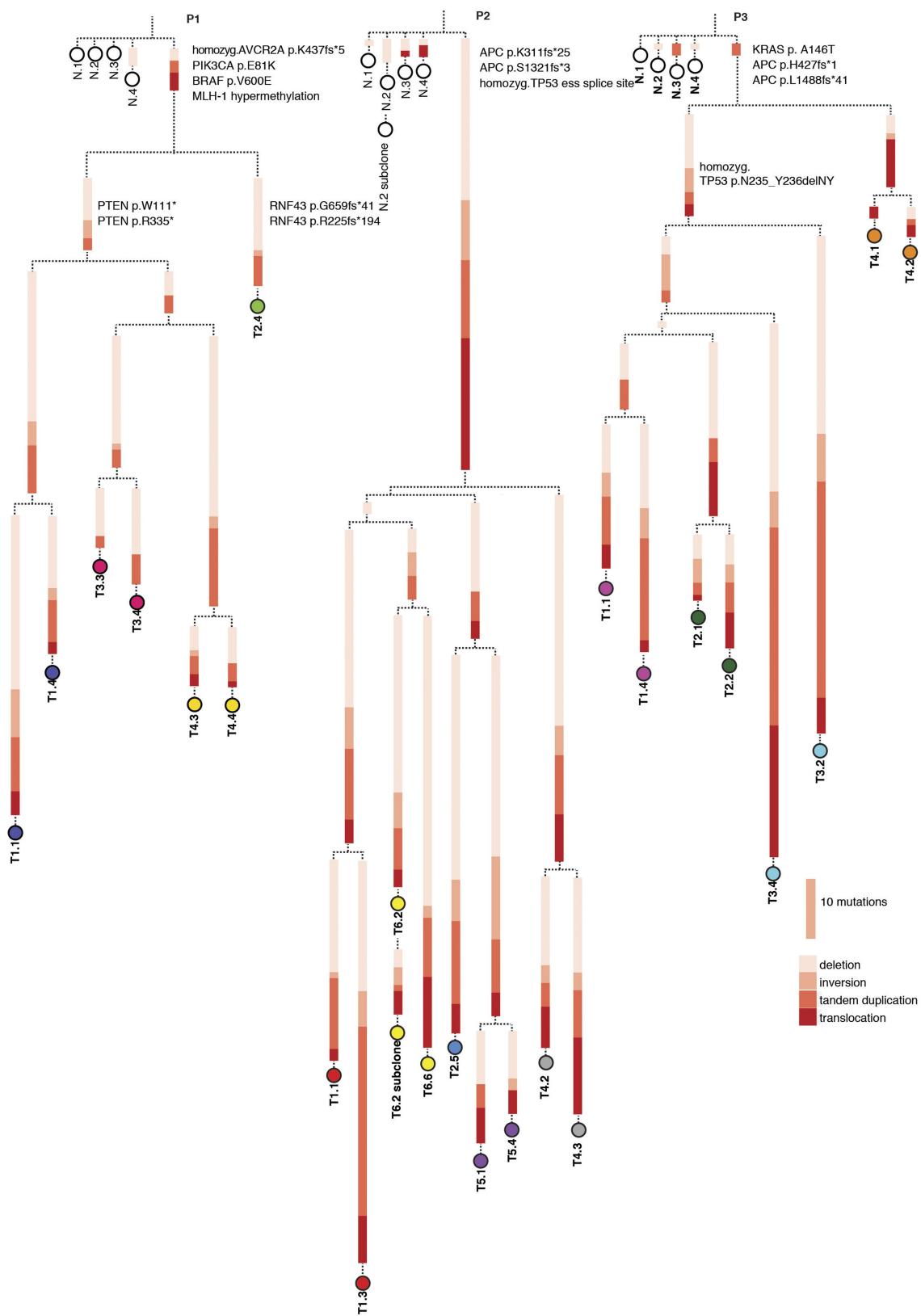
Extended Data Fig. 5 | Phylogenetic trees for rearrangements.
Phylogenies for three individuals with branch lengths representing rearrangement numbers, further subdivided into deletions, inversions, tandem duplications and translocations.
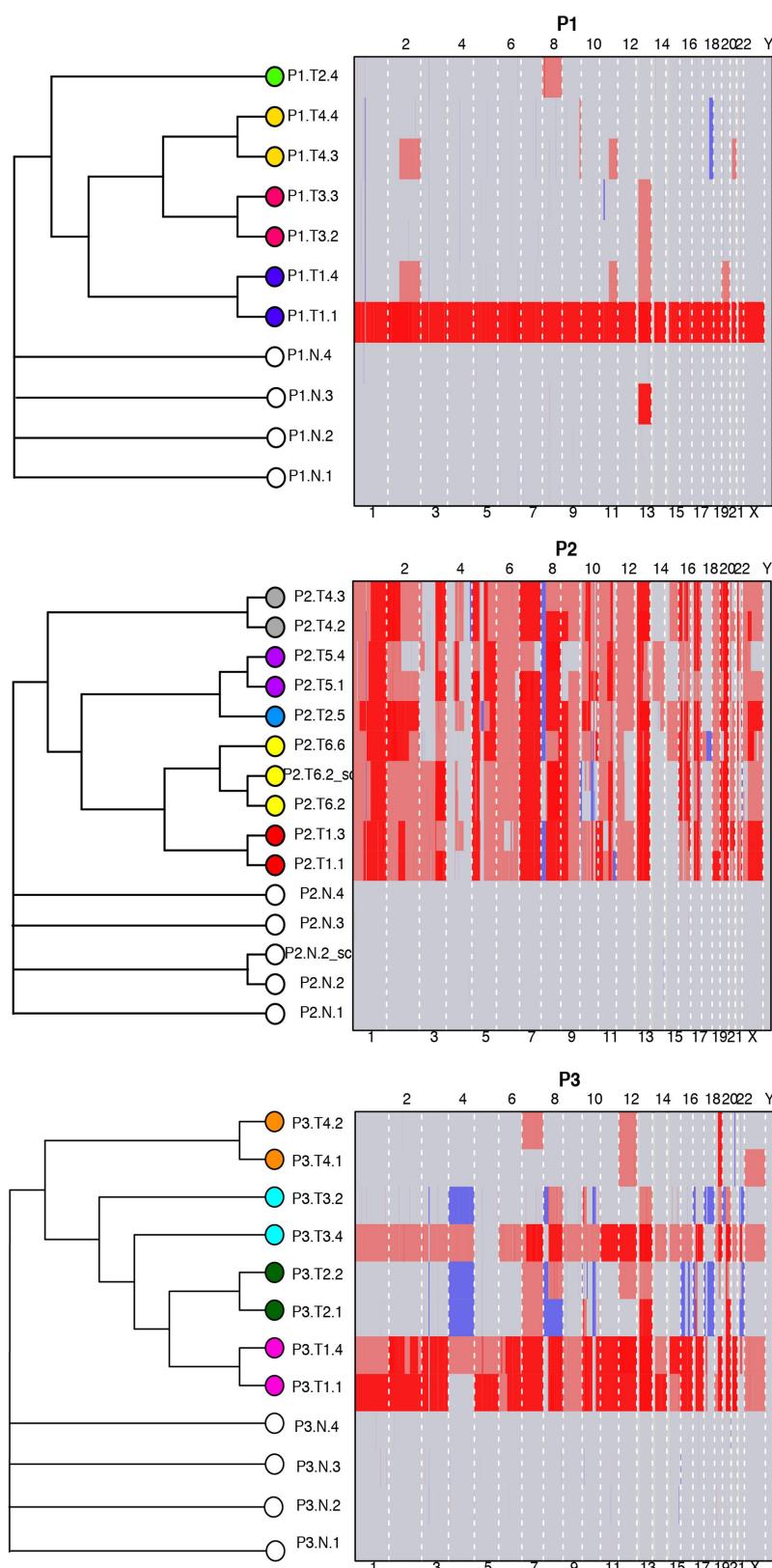
Extended Data Fig. 6 | Copy number analysis. Copy number profiles of all clones that have been WGS analysed, displayed as a heatmap (amplification in red, loss in blue). The structures of the phylogenetic trees are displayed on the left; branch lengths are not scaled.
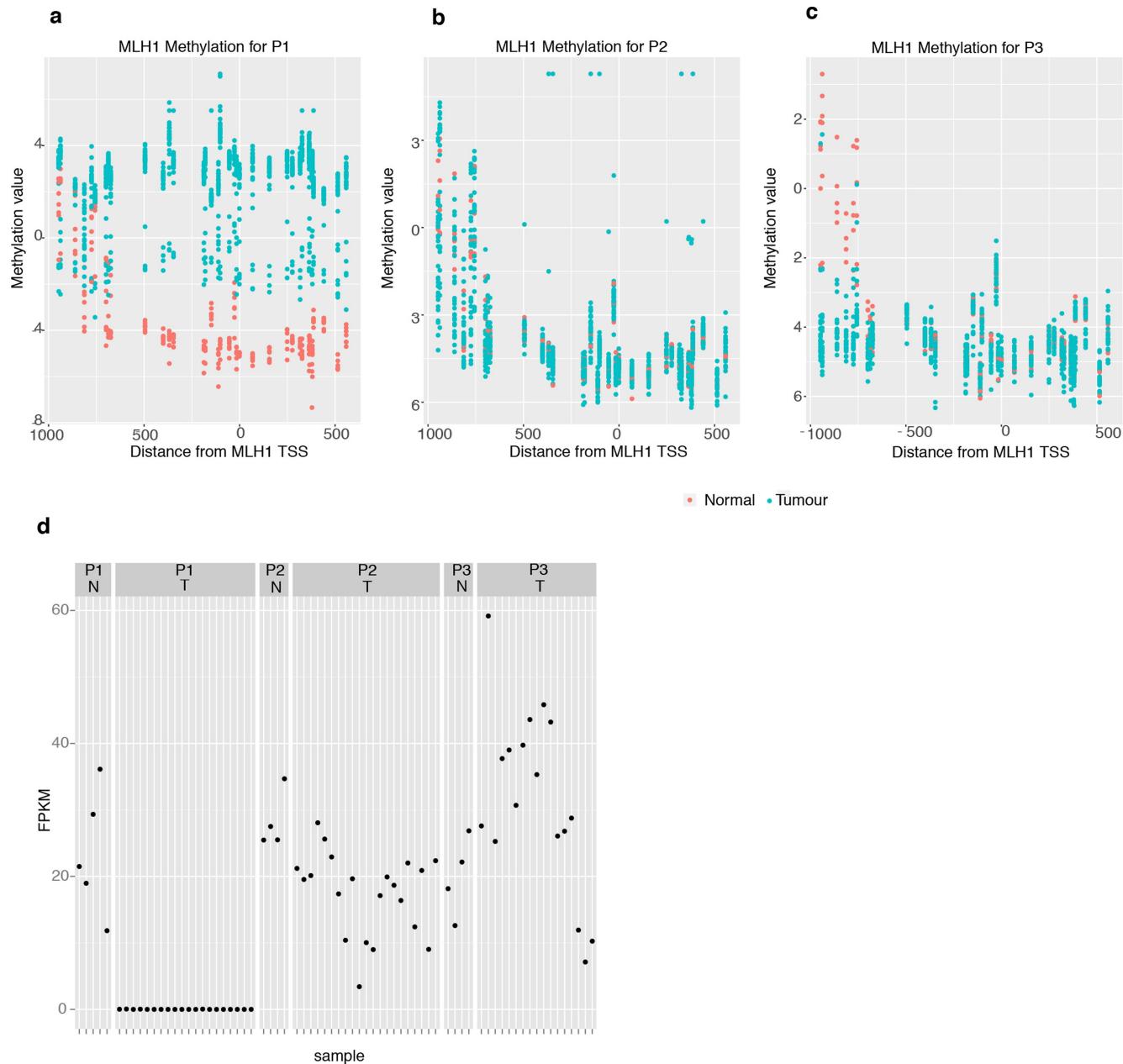
Extended Data Fig. 7 | *MLH1* hypermethylation in P1. a–c, Methylation pattern of the *MLH1* gene for tumour and normal clones for three individuals, showing hypermethylation in proximity to the transcription start site (TSS) for P1 tumour clones compared to normal clones. **d,** Expression of *MLH1* in all clones; *MLH1* transcript could not be detected in tumour clones from P1.
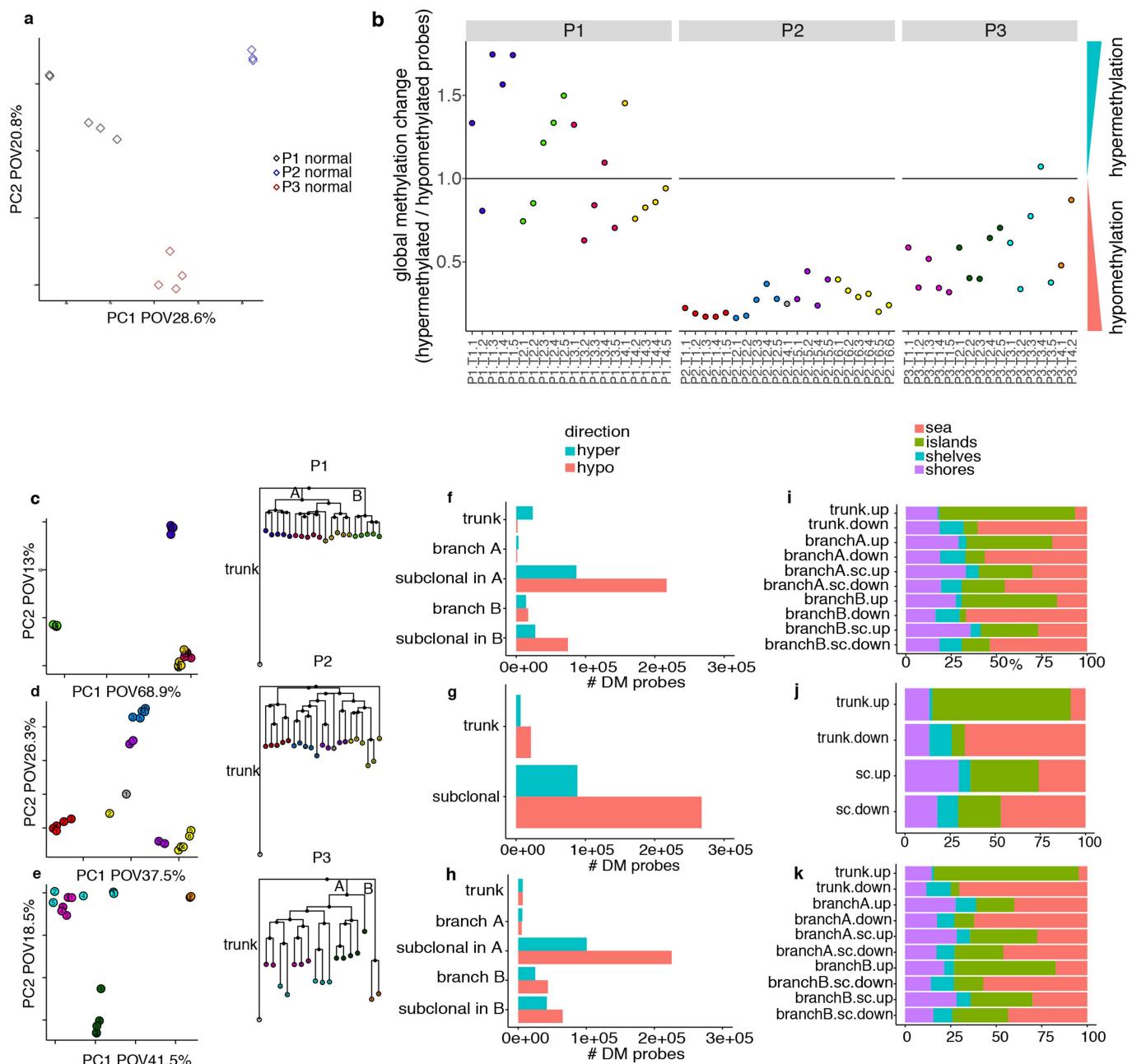
Extended Data Fig. 8 | Methylation analysis. a, Clustering of methylation data by PCA showing normal-derived organoids from three individuals ($n = 12$ biologically independent samples). b, Global methylation change in each tumour clone, expressed as the ratio of hypermethylated probes to hypomethylated probes. Hyper- and hypomethylation are assessed by comparing to the baseline methylation levels in the normal-derived clones (indicated with line at $y = 1$). c–e, Left, clustering of methylation data by PCA of tumour organoids from each individual, displaying the first two principal components. Clones from different segments are shown in different colours as in Extended Data Fig. 2. Right, phylogenetic trees based on expression data (as in Fig. 3b) with the main branches used for our methylation analysis indicated. c, P1, $n = 20$ biologically independent samples. d, P2, $n = 21$ biologically independent samples. e, P3, $n = 17$ biologically independent samples. f–h, Direction of methylation changes during tumour development. Methylation changes were assigned to either the branch of the tumour or the main subclonal branches (indicated in the phylogenetic trees in e). i–k, Relative proportion of probes in CpG islands, shores, shelves and seas that were differentially methylated in different branches (Supplementary Notes section 6).
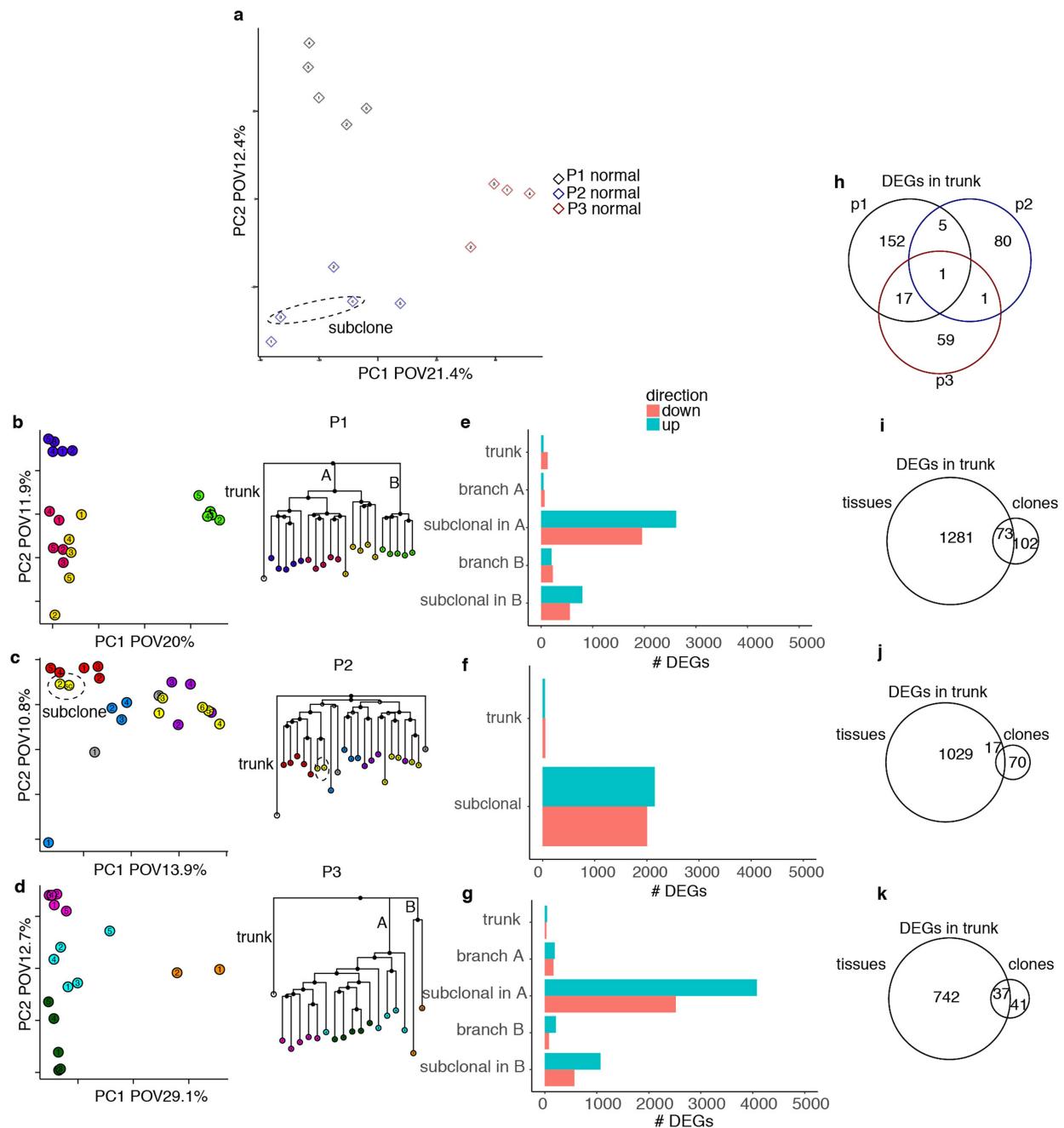
Extended Data Fig. 9 | Expression analysis. a, PCA based on expression pattern of normal organoids from each individual, displaying the first two principal components ($n = 13$). A subclone and its ancestral clone are circled. **b–d**, Left, PCA of tumour clones from each individual. Clones derived from different segments are shown in different colours as in Figs. 2–4. A subclone derived from a tumour clone from P2 and its ancestor clone are circled. Right, phylogenetic trees based on expression data (as in Fig. 4b) with the main branches used for our expression analysis indicated. **b**, P1, $n = 20$ biologically independent samples. **c**, P2, $n = 22$ biologically independent samples. **d**, P3, $n = 17$ biologically independent

samples. **e–g**, Global analysis of expression changes attributed to the trunk of the tree, the main branches or subclonal variation. **h**, Venn diagram displaying the differentially expressed genes that were attributed to the trunk of each tumour. Differentially expressed genes determined by a likelihood ratio test using a negative binomial generalized linear model fit (FDR < 0.05). **i–k**, Comparison of differentially expressed genes identified in the organoid clones of each patient versus the original tissue sections. Only genes that were significantly altered in all clones or all biopsies from each individual are considered.
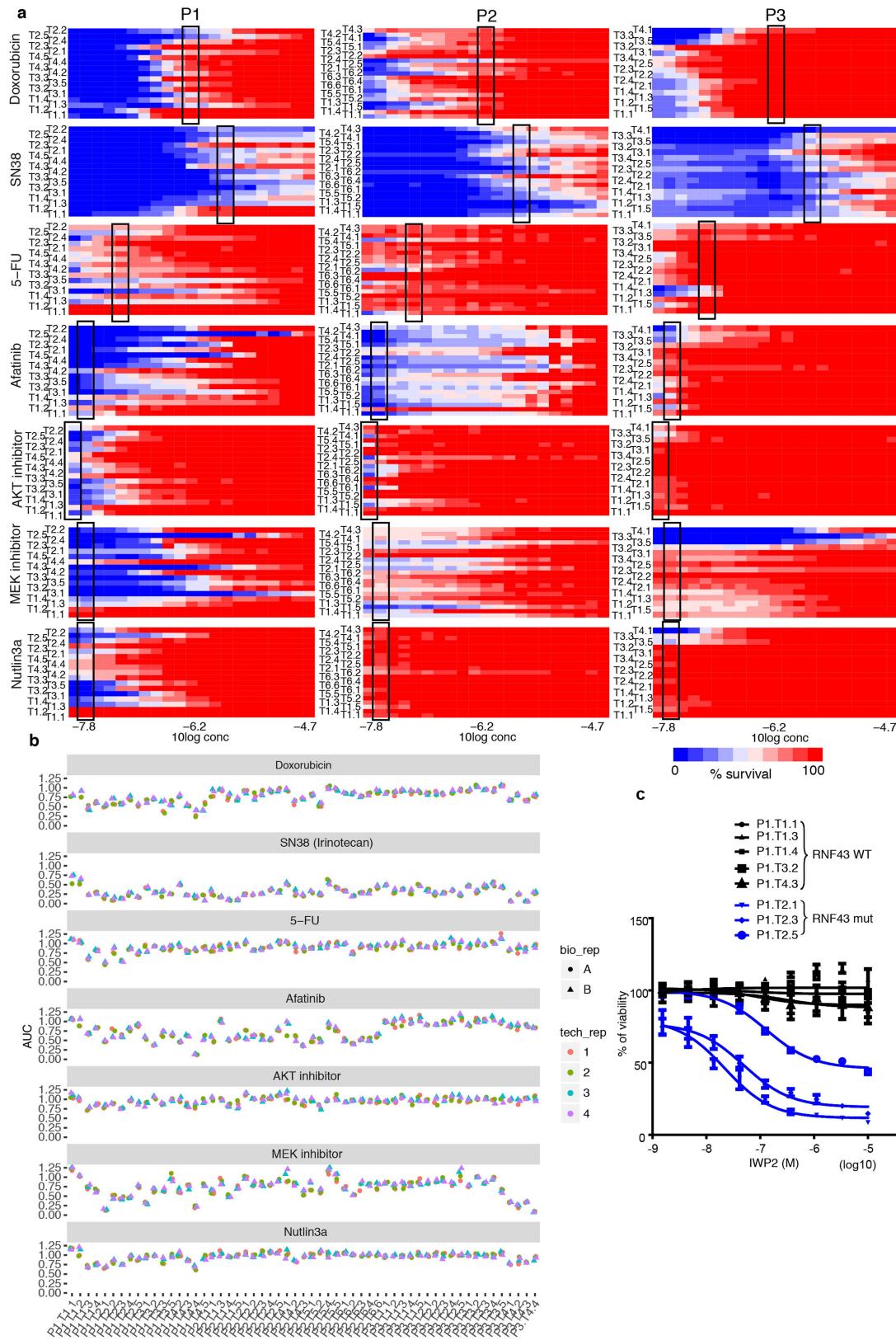
Extended Data Fig. 10 | Drug response data. Dose response data for seven drugs, tested on organoid clones from three individuals. Twenty-one concentrations were tested for each drug, ranging from 14.7 nm to 20 μM. Mean survival from two duplicate experiments is displayed in a heatmap. The concentration displayed in Fig. 4 is outlined with a black box in each panel. **b**, Reproducibility of drug response data. Each measurement was performed twice (technical replicate) and each experiment was performed in duplicate (biological replicate). For each biological or technical replicate the area under the curve (AUC) is shown. **c**, Dose–response curves after 6 days of treatment with IWP2 (Wnt secretion inhibitor) for clonal tumour organoids derived from P1. *RNF43* mutant clones are responsive, whereas *RNF43* wild-type (WT) clones are resistant. Data points and error bars represent the mean and s.d. of four independent technical replicates from two independent experiments.

ARTICLE

Identification of the tumour transition states occurring during EMT

Ievgenia Pastushenko[1], Audrey Brisebarre[1], Alejandro Sifrim[2,3], Marco Fioramonti[1], Tatiana Revenco[1], Soufiane Boumahdi[1], Alexandra Van Keymeulen[1], Daniel Brown[2,4], Virginie Moers[1], Sophie Lemaire[1], Sarah De Clercq[5], Esmeralda Minguijón[5], Cédric Balsat[6], Youri Sokolow[7], Christine Dubois[1], Florian De Cock[1], Samuel Scozzaro[1], Federico Sopena[8], Angel Lanas[9], Nicky D'Haene[5], Isabelle Salmon[5,6], Jean-Christophe Marine[4,10], Thierry Voet[2,3], Panagiota A. Sotiropoulou[1,12] & Cédric Blanpain[1,11,12]*

In cancer, the epithelial-to-mesenchymal transition (EMT) is associated with tumour stemness, metastasis and resistance to therapy. It has recently been proposed that, rather than being a binary process, EMT occurs through distinct intermediate states. However, there is no direct in vivo evidence for this idea. Here we screen a large panel of cell surface markers in skin and mammary primary tumours, and identify the existence of multiple tumour subpopulations associated with different EMT stages: from epithelial to completely mesenchymal states, passing through intermediate hybrid states. Although all EMT subpopulations presented similar tumour-propagating cell capacity, they displayed differences in cellular plasticity, invasiveness and metastatic potential. Their transcriptional and epigenetic landscapes identify the underlying gene regulatory networks, transcription factors and signalling pathways that control these different EMT transition states. Finally, these tumour subpopulations are localized in different niches that differentially regulate EMT transition states.

EMT is a cellular process in which cells lose their epithelial characteristics and acquire mesenchymal features, which enable them to migrate more efficiently and invade the underlying mesenchyme. EMT is essential for gastrulation, somitogenesis and neural crest delamination during embryonic development and has been associated with various diseases. In cancer, EMT is associated with tumorigenesis, invasion, metastasis, tumour stemness and resistance to therapy[1,2]. Although EMT has traditionally been viewed as a binary switch, some in vitro data (mainly co-expression of epithelial and mesenchymal markers within the same cells) have indicated that EMT may proceed in a stepwise manner through the generation of subpopulations that represent different intermediate states between the epithelial and mesenchymal states[3–7]. However, it remains unclear whether EMT proceeds through these intermediate states in vivo, and if so how many intermediate steps exist, how plastic and reversible these intermediate states are, which mechanisms regulate the transition from one state to another and what the implications of these different EMT states are for tumour progression, stemness and metastasis[1,2].

Different tumour EMT transition states

To determine whether EMT in vivo occurs through a succession of different intermediate states, we used a genetic mouse model of skin squamous cell carcinoma (SCC) mediated by the conditional expression of $KRas^{G12D}$ and $p53$ deletion ($p53^{KO}$) in hair follicles. This model generates skin tumours that undergo spontaneous EMT, containing epithelial YFP^+Epcam^+ and mesenchymal-like YFP^+Epcam^- tumour cells (TCs)[8]. Using flow cytometry (fluorescence-activated cell sorting, FACS), we screened cells from these tumours for a large panel of cell surface markers and assessed whether these markers were

heterogeneously expressed in YFP^+Epcam^+ or YFP^+Epcam^- populations (Fig. 1a). The YFP^+Epcam^+ TCs were relatively homogenous, with only four markers being heterogeneously expressed (Fig. 1b). By contrast, half of the markers were heterogeneously expressed in YFP^+Epcam^- TCs (Fig. 1b, Extended Data Fig. 1a, b), suggesting that EMT is associated with important cellular heterogeneity. The markers that were most frequently heterogeneously expressed during EMT included CD61 (also known as Itgb3), CD51 (also known as Itgav) and CD106 (also known as Vcam1; Extended Data Fig. 1c), which mark subpopulations of TCs associated with tumour stemness, EMT or metastasis initiation in other tumour models[7,9–12]. Other markers were not as frequently heterogeneously expressed when analysed in a larger cohort of tumours (Extended Data Fig. 1c). Combinatorial multicolour FACS analysis revealed that CD106, CD61, and CD51 discriminated six distinct populations within YFP^+Epcam^- TCs in most (75%) mixed tumours (Fig. 1c, e, h, Extended Data Fig. 2). About 10% of mixed tumours with a high proportion of Epcam+ cells presented only triple-negative ($Epcam^-$ $CD51^-CD61^-CD106^-$) and $Epcam^-$ $CD51^+CD61^+$ populations, whereas highly mesenchymal tumours with only $Epcam^-$ TCs contained $CD51^+$, $CD51^+CD61^+$, $CD106^+CD51^+$ and $CD106^+CD51^+CD61^+$ triple-positive tumour subpopulations with almost no triple-negative and $CD106^+$ subpopulations (Fig. 1c–g).

To define whether these different tumour populations correspond to distinct EMT transition states, we isolated the subpopulations by FACS and performed immunostaining on cytospin with epithelial (keratin 14, K14) and mesenchymal (vimentin) markers. Notably, loss of Epcam expression coincided with a gain in vimentin expression in all TCs, consistent with the first switch to the mesenchymal state (Fig. 1i, j). However,

Fig. 1 | Identification of the tumour transition states occurring during EMT in vivo. a, Strategy for marker screening in hair follicle-derived skin SCC model presenting spontaneous EMT. **b**, Percentage of homogeneously and heterogeneously expressed markers in YFP⁺Epcam⁺ and YFP⁺Epcam⁻ TCs. Two-tailed Fisher's exact test. **c**, Percentage of Epcam⁺ cells in the four groups of tumours that differ by the frequencies of the different subpopulations. **d–g**, Distributions of subpopulations in tumours with pattern 1 (**d**; mixed tumours containing Epcam⁺ and Epcam⁻ TCs), pattern 2 (**e**; differentiated tumours with a high percentage of Epcam⁺ TCs), pattern 3 (**f**) and pattern 4 (**g**; mesenchymal Epcam⁻ TCs) ($n = 12$ mice, mean ± s.e.m.). TN, triple-negative; TP, triple-positive. **h**, FACS profile showing the six CD106/CD51/CD61 subpopulations in the most frequent pattern (pattern 1). **i**, Co-immunostaining of keratin 14 (K14) and vimentin (Vim) in cystospin of FACS-isolated tumour cells. Scale bars, 20 μm; $n = 6$. **j**, Cytospin counts of TCs based on K14 and vimentin expression (average of 90 cells per condition and tumour). **k**, PCA of scRNA-seq of Epcam⁺ and Epcam⁻ TCs. Dots represent single cells; colour scale represents the normalized expression of each gene ($n = 66$ Epcam⁺ and $n = 277$ Epcam⁻ cells from one tumour). **l**, Proportion of single cells expressing Epcam, CD51, CD61 and CD106 markers. **m**, PCA plot coloured by expression of the markers used to define EMT subpopulations by FACS. Grey circle, epithelial tumour cells; blue, hybrid cells; yellow, early EMT cells; red, late EMT cells. **n**, Cumulative density of each marker combination across PC1, which correspond to the degree of EMT (**j**).

some Epcam⁻ tumour subpopulations (triple-negative, CD106⁺ and CD51⁺) continued to express K14 and vimentin, whereas other Epcam⁻ subpopulations (CD51⁺CD61⁺ and triple-positive) were essentially K14⁻vimentin⁺, with rare TCs expressing low levels of K14 (Fig. 1i, j). These data indicate that the subpopulations identified during spontaneous EMT of primary skin tumours correspond to different tumour populations with different degrees of EMT, with some subpopulations corresponding to the hybrid tumour phenotypes with epithelial and mesenchymal features that have been described in cancer cell lines in vitro[3–7].

To further assess cellular heterogeneity during spontaneous EMT, we performed single cell RNA sequencing (scRNA-seq) of FACS-isolated YFP⁺Epcam⁺ and YFP⁺Epcam⁻ TCs. Dimensionality reduction using principal component analysis (PCA) revealed that the first principal component (PC1), which explained 21% of the variability, could be attributed to EMT state (Fig. 1k). Our scRNA-seq data confirmed that Epcam⁻ subpopulations showed greater transcriptional heterogeneity than Epcam⁺ subpopulations (Fig. 1k). The expression of epithelial and mesenchymal markers at the single-cell level confirmed the progressive acquisition of EMT features with epithelial, mesenchymal and hybrid states (Extended Data Fig. 3) and the distribution of CD51, CD61 and CD106 markers correlated with the degree of EMT along PC1, confirming the EMT gradient found across the different tumour subpopulations (Fig. 1l–n).

To assess whether these different subpopulations of EMT TCs reflect a more general mechanism occurring during EMT, we assessed the expression of these cell surface markers in metaplastic-like mammary tumours arising from oncogenic *Pik3ca* expression and *p53* deletion and in *MMTV-PyMT* mammary luminal tumours, which have been reported to present EMT features[13–16]. Notably, a subset of mammary metaplastic-like and *MMTV-PyMT* luminal tumours also contained Epcam⁺ and Epcam⁻ TCs that could be subdivided into the same six subpopulations as found in *KRas^{G12D}/p53^{KO}* skin tumours (Extended Data Figs. 4a–d, 5a–c). Immunostaining on cytospin and real-time PCR with reverse transcriptase (RT–PCR) showed that the subpopulations isolated from the mammary tumours presented different degrees of EMT, similar to those identified in skin SCCs (Extended Data Figs. 4e, f, h, 5d, e, g), demonstrating that the different EMT transition states identified here represent a conserved mechanism during EMT.

To investigate whether these EMT transition states exist in human cancers, we assessed the expression of epithelial and mesenchymal markers in tumours derived from xenotransplantation (PDX) of poorly differentiated human breast cancers and SCCs. After several passages in immunodeficient mice, human stroma is entirely replaced by mouse cells[17], making it possible to differentiate human TCs that underwent EMT and lost the expression of epithelial markers from mouse stroma using an antibody against human antigen (Ku-80). We detected areas expressing only epithelial markers, areas co-expressing epithelial and mesenchymal markers and areas expressing exclusively mesenchymal markers in poorly differentiated breast cancer, lung and oesophageal SCCs (Extended Data Fig. 6). These data demonstrate that EMT in human cancers is associated with different transition states including hybrid states, as was suggested by scRNA-seq of human SCCs[18].

Stemness and plasticity of EMT states

EMT has been associated with cancer stemness, characterized by an increase in tumour-propagating cell (TPC) frequency[1,8,19,20]. As previously described[8], Epcam⁻ TCs contained five times as many TPCs as did Epcam⁺ TCs (Fig. 2a). Notably, all EMT subpopulations presented similar TPC frequencies (Fig. 2a). These data show that the earliest EMT state already exhibits increased TPC frequency, and tumour stemness does not increase further in later transition states. Whereas Epcam⁺ TCs showed higher proliferation than Epcam⁻ TCs, there was no difference in proliferation rate between the different EMT subpopulations (Fig. 2b). Thus, TPC frequency is inversely correlated to in vivo proliferation.

Cancer cells have been shown to be plastic in transplantation assays, with different tumour subpopulations able to transit back and forth between different states and to recapitulate primary tumour heterogeneity[21]. To determine whether the different subpopulations identified during EMT are similarly plastic, we analysed the tumour phenotypes of secondary tumours. As previously described[8], secondary tumours arising from the transplantation of Epcam⁻ TCs comprise only Epcam⁻ TCs (Fig. 2c). Although all EMT subpopulations presented a certain degree of plasticity, at early time points following transplantation (3–4 weeks), the triple-negative subpopulation was relatively primed towards the epithelial phenotype and preferentially gave rise to
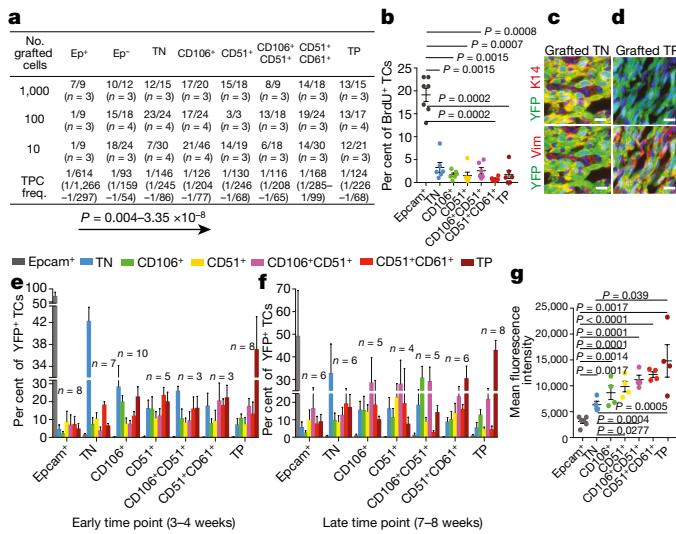
Fig. 2 | EMT transition states present similar TPC capacity but exhibit differing plasticity. a, Frequency of secondary tumours observed upon transplantation of limiting dilution of TC subpopulations and the estimation of TPC frequency (χ^2 test). **b**, FACS quantification of BrdU incorporation in TC subpopulations. Mean ± s.e.m., two-tailed *t*-test. **c, d**, Immunofluorescence of YFP and K14 or vimentin in secondary tumours arising after subcutaneous transplantation of TN (**c**) or TP (**d**) cells ($n = 3$; scale bars, 20 μm). **e, f**, Proportion of each subpopulation in secondary tumours 3–4 weeks (**e**) and 7–8 weeks (**f**) after subcutaneous transplantation of the different subpopulations (mean + s.e.m.). **f**, Invasion ability of different subpopulations measured by in vitro cell invasion assay. Mean fluorescence intensity correlates with the number of cells that migrate through the ECM. Two-tailed *t*-test, mean ± s.e.m., $n = 3$.

triple-negative TCs. By contrast, the most mesenchymal triple-positive subpopulation preferentially gave rise to triple-positive TCs (Fig. 2c–e). The other EMT subpopulations were plastic, giving rise to each other in similar proportions (Fig. 2c). Notably, at later time points (7–8 weeks), the triple-negative subpopulation gave rise to different EMT subpopulations, whereas the triple-positive subpopulation gave rise preferentially to triple-positive cells (Fig. 2f). The different tumour subpopulations had different invasive capacities, which increased with the degree of EMT (Fig. 2f). Together, these data show that the different EMT tumour subpopulations are functionally distinct and have different clonogenic potentials, invasive properties, differentiation abilities and plasticity phenotypes.

Metastasis of EMT transition states

It remains unclear whether EMT is associated with increased metastatic potential[1,14,22–25]. We have previously shown that Epcam[−] cells are more able than Epcam[+] TCs to metastasize to the lung[8]. To determine whether some EMT subpopulations showed increased ability to undergo vascular extravasation, lung colonization and metastasis, we injected these subpopulations of TCs intravenously into mice and assessed their ability to give rise to lung metastasis. Triple-negative and CD106[+] hybrid subpopulations gave rise to substantially more metastases than did other EMT subpopulations (Fig. 3a). The increase in the metastatic potential of the hybrid states was also observed in mammary metaplastic-like and *MMTV-PyMT* luminal tumours (Extended Data Figs. 4g, 5f). Whereas CD106 has been shown to be associated with increased metastatic potential in human cancer cell lines[9,11], our data show that not all subpopulations of cells expressing CD106 are equally metastatic; their metastatic potential correlated more with the hybrid EMT state than the level of CD106 expression (Fig. 3a). As intravenous injection does not allow us to investigate the ability of cells to leave primary tumours and intravasate into blood vessels, we assessed the presence of YFP[+] circulating tumour cells (CTCs) in the blood of *Lgr5Cre/ERKras*[G12D]*/p53*[KO] mice. The majority of CTCs in this mouse model underwent EMT, and the majority
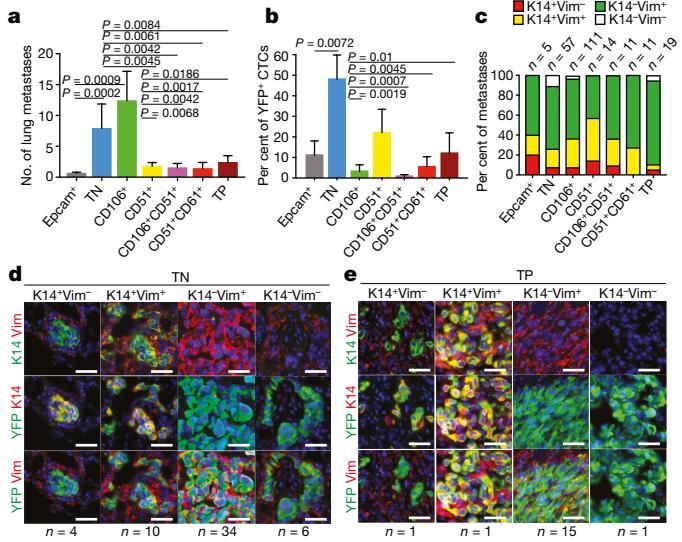

Fig. 3 | Different EMT transition states present distinct metastatic potential. a, Number of lung metastases arising from the injection of 500 YFP[+] TCs with different degrees of EMT ($n = 8$). **b**, Phenotype of YFP[+] CTCs based on expression of Epcam, CD106, CD51 and CD61 ($n = 10$). **a, b**, Two-tailed Mann–Whitney test, mean + s.e.m. **c**, Classification of metastasis based on K14 and vimentin expression (mean percentage). **d, e**, Co-immunostaining of YFP and K14 or vimentin in metastases arising from intravenously injected TN (**d**) or TP (**e**) TCs. Scale bars, 50 μm.

of Epcam[−] CTCs were triple-negative (Fig. 3b), further supporting the notion that the hybrid EMT state is associated with increased metastatic potential.

Mesenchymal-to-epithelial transition (MET) has been associated with increased metastatic potential in several mouse cancer models that are induced to undergo EMT by forced overexpression of EMT transcription factors such as Twist1 and Prrx1[26,27]. Whereas subcutaneously injected Epcam[−] cells did not give rise to Epcam[+] TCs in secondary tumours, all Epcam[−] subpopulations were able to revert back to Epcam[+] TCs when colonizing the lung (Fig. 3c–e), showing that even the most extreme EMT states are not irreversibly locked into the mesenchymal state and can undergo MET within the lung microenvironment. However, the number of metastases did not correlate with the greater ability of these populations to undergo MET (Fig. 3c), suggesting that other mechanisms beside MET contribute to the higher metastatic potential of these hybrid populations.

Molecular characterization of EMT states

To unravel the molecular mechanisms associated with the different EMT tumour subpopulations, we performed RNA-seq and assay for transposase-accessible chromatin using sequencing (ATAC-seq) of the different FACS-isolated tumour subpopulations to define their transcriptional and chromatin landscapes. RNA-seq revealed that the level of expression of different epithelial markers, including cytoskeletal elements, adhesion molecules and epithelial transcription factors, progressively decreased from Epcam[+] to triple-negative, CD106[+] and CD51[+] cells, and then remained very low in CD106[+]CD51[+], CD51[+]CD61[+] and triple-positive subpopulations (Fig. 4a). By contrast, some mesenchymal markers increased strongly with the loss of Epcam expression (in triple-negative cells), and remained stable, whereas other EMT markers such as *Fn1*, *Prrx1*, *Col3a1* and *Lox* presented a clear gradient of expression with maximal expression in triple-positive cells (Fig. 4b). This was further illustrated by unsupervised clustering analysis of the RNA-seq data (Extended Data Fig. 7a).

To further delineate the underlying molecular differences between the different EMT states, we performed unsupervised clustering analysis of the scRNA-seq data. Using SC3 clustering, we identified different tumour populations including epithelial, hybrid and mesenchymal populations (Extended Data Fig. 8a, b). Pseudotime ordering using

Fig. 5 | Different EMT transition states are localized in different niches. a–d, Mosaic images of immunostaining for YFP (a), K14 (b), vimentin (c) and CD61 (d) in a representative mixed SCC (n = 5). e, Definition of four areas based on the expression of K14, vimentin and CD61. Red, area expressing vimentin; green, area expressing K14; yellow, area expressing vimentin and K14; white outline, area expressing CD61. f–i, Expression of K14, vimentin, YFP and CD61 in area 1 (f), area 2 (g), area 3 (h) and area 4 (i). j–m, Expression of YFP and CD45, CD68, CD31 or Lyve1 in area 1 (j), area 2 (k), area 3 (l) and area 4 (m). Scale bars, 50 μm. n, mRNA expression of chemokine, pro-inflammatory and pro-angiogenic genes by TCs (mean + s.e.m.). o, p, Percentage of F4/80$^+$CD11b$^+$ macrophages in peripheral blood (o) and number of tumours (p) in control (n = 3) and anti-Csf1r/Ccl2-treated (n = 3) mice. q, r, Proportion of Epcam$^+$ TCs (q) and distribution of different EMT subpopulations (r) in control and anti-Csf1r/Ccl2-treated mice (n = 26 tumours from n = 3 control mice and n = 15 tumours from n = 3 anti-Csf1r/Ccl2-treated mice). o–r, Two-tailed t-test, mean ± s.e.m.

Smad2 promote the transition from the Epcam$^+$ and triple-negative states towards more mesenchymal states.

Different niches of the EMT states

To determine whether these different tumour subpopulations are localized in particular areas within tumours, we performed immunostaining with markers characteristic of the distinct tumour subpopulations. YFP expression marked TCs irrespective of their degree of EMT. Mixed tumours could be subdivided into distinct areas corresponding to different parts of the tumour, with differing degrees of EMT based on the expression of K14, vimentin and CD61 (Fig. 5a–e). Area 1 corresponded to the pure epithelial part of the tumours. In area 2, the TCs were more elongated, although they were still cohesive. K14 and vimentin were co-expressed, marking the hybrid epithelial and mesenchymal states of the tumours. In area 3, TCs were further elongated with a more fibroblast-like appearance; some TCs did not show cell–cell

adhesion, expressed vimentin and had lost K14 expression. Most of the cells in this area expressed CD61. In area 4, YFP$^+$ TCs were even more fibroblastic, did not form cell–cell junctions, did not express K14 and uniformly expressed vimentin and CD61 (Fig. 5f–i). These data reveal that the different EMT tumour subpopulations are spatially organized in specific regions rather than being randomly distributed throughout tumours.

To determine whether the TC subpopulations described here are localized in particular microenvironments, we investigated whether they were associated with specific stromal populations. In area 1, endothelial cells (CD31$^+$), inflammatory cells (CD45$^+$) and cancer-associated fibroblasts (YFP$^-$vimentin$^+$) were relatively rare and surrounded the epithelial part of the tumour. By contrast, the composition and the localization of stromal cells changed markedly as TCs progressed towards EMT. Cells that were in close contact with YFP$^+$ EMT TCs showed a large increase in the density of CD45$^+$ immune cells, particularly monocytes and macrophages expressing CD68, and an increase in the density of endothelial and lymphatic cells (Fig. 5j–m). These different areas—composed of epithelial, hybrid and mesenchymal cells with increased vascularization and inflammatory cells—were also observed in Pik3ca/p53^{KO} and MMTV-PyMT mammary tumours (Extended Data Figs. 4i–p, 5h–m). Notably, the immune infiltration density correlated with the expression of chemokines and other pro-inflammatory and pro-angiogenic molecules by TCs (Fig. 5n), suggesting that TCs attract and regulate the stromal cells that comprise their niche and contribute directly to the spatial organization of tumour subpopulations.

To determine the functional relevance of the macrophage infiltration in regulating EMT progression, we assessed the effect of macrophage depletion on the proportions of EMT tumour subpopulations. Administration of anti-Csf1r and anti-Ccl2 blocking antibodies during tumorigenesis efficiently depleted circulating and tumour-infiltrating macrophages and decreased tumour formation; in addition, it increased the proportion of Epcam$^+$ and triple-negative TCs, and decreased the proportion of triple-positive TCs (Fig. 5o–r, Extended Data Fig. 10). These results demonstrate the functional importance of the tumour microenvironment and macrophage infiltration in regulating the transition between EMT states.

Discussion

Our study demonstrates that spontaneous EMT in primary TCs in vivo proceeds through distinct intermediate states with different invasive, metastatic and differentiation characteristics. TCs with hybrid epithelial and mesenchymal phenotypes are more efficient in reaching the circulation, colonizing the lungs and forming metastases. Transcriptional, chromatin and scRNA profiling demonstrate that these different states have distinct cellular properties, chromatin landscapes and gene expression signatures that are regulated by common and distinct transcription factors and signalling pathways. In addition, these different EMT states are localized in different microenvironments and in contact with different stromal cells. The most mesenchymal subpopulations are localized close to endothelial and inflammatory cells. These TCs secrete high levels of chemokines and proteins that attract immune cells and promote angiogenesis, thereby organizing the formation of their own inflammatory and highly vascularized niche; this in turn further promotes EMT in these subpopulations. Our results identify different EMT transition states in vivo that are associated with different tumour functions, and have important implications for our understanding of tumour heterogeneity, growth, invasion, metastasis and resistance to therapy. The approaches used here to unravel EMT-related tumour heterogeneity can be used to define cellular heterogeneity in other biological systems.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0040-3.

1. Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
2. Puisieux, A. & Brabletz, T. & Caramel, J. Oncogenic roles of EMT-inducing transcription factors. *Nat. Cell Biol.* **16**, 488–494 (2014).
3. Huang, R. Y. et al. An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell Death Dis.* **4**, e915 (2013).
4. Zhang, J. et al. TGF-β-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci. Signal.* **7**, ra91 (2014).
5. Hong, T. et al. An Ovol2-Zeb1 mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states. *PLoS Comput. Biol.* **11**, e1004569 (2015).
6. Jolly, M. K. et al. Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget* **7**, 27067–27084 (2016).
7. Bierie, B. et al. Integrin-β4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells. *Proc. Natl Acad. Sci. USA* **114**, E2337–E2346 (2017).
8. Latil, M. et al. Cell-type-specific chromatin states differentially prime squamous cell carcinoma tumor-initiating cells for epithelial to mesenchymal transition. *Cell Stem Cell* **20**, 191–204 (2017).
9. Chen, Q., Zhang, X. H. & Massagué, J. Macrophage binding to receptor VCAM-1 transmits survival signals in breast cancer cells that invade the lungs. *Cancer Cell* **20**, 538–549 (2011).
10. Seguin, L. et al. An integrin β3–KRAS–RalB complex drives tumour stemness and resistance to EGFR inhibition. *Nat. Cell Biol.* **16**, 457–468 (2014).
11. Slack-Davis, J. K., Atkins, K. A., Harrer, C., Hershey, E. D. & Conaway, M. Vascular cell adhesion molecule-1 is a regulator of ovarian cancer peritoneal metastasis. *Cancer Res.* **69**, 1469–1476 (2009).
12. Wang, J. et al. CD51 correlates with the TGF-beta pathway and is a functional marker for colorectal cancer stem cells. *Oncogene* **36**, 1351–1363 (2017).
13. Van Keymeulen, A. et al. Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* **525**, 119–123 (2015).
14. Ye, X. et al. Upholding a role for EMT in breast cancer metastasis. *Nature* **547**, E1–E3 (2017).
15. Ye, X. et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* **525**, 256–260 (2015).
16. DelPozo Martin, Y. et al. Mesenchymal cancer cell-stroma crosstalk promotes niche activation, epithelial reversion, and metastatic colonization. *Cell Reports* **13**, 2456–2469 (2015).
17. Hidalgo, M. et al. Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov.* **4**, 998–1013 (2014).
18. Puram, S. V. et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171**, 1611–1624 (2017).
19. Mani, S. A. et al. The epithelial–mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
20. Ye, X. & Weinberg, R. A. Epithelial-mesenchymal plasticity: a central regulator of cancer progression. *Trends Cell Biol.* **25**, 675–686 (2015).
21. Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328–337 (2013).
22. Diepenbruck, M. & Christofori, G. Epithelial–mesenchymal transition (EMT) and metastasis: yes, no, maybe? *Curr. Opin. Cell Biol.* **43**, 7–13 (2016).
23. Zheng, X. et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525–530 (2015).
24. Fischer, K. R. et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472–476 (2015).
25. Aiello, N. M. et al. Upholding a role for EMT in pancreatic cancer metastasis. *Nature* **547**, E7–E8 (2017).
26. Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S. & Yang, J. Spatiotemporal regulation of epithelial–mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **22**, 725–736 (2012).
27. Ocaña, O. H. et al. Metastatic colonization requires the repression of the epithelial–mesenchymal transition inducer Prrx1. *Cancer Cell* **22**, 709–724 (2012).
28. Siegel, P. M. & Massagué, J. Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer. *Nat. Rev. Cancer* **3**, 807–821 (2003).

Author contributions I.P. and C.Bla. designed the experiments and performed data analysis. I.P. performed most of the biological experiments. P.A.S. performed part of the FACS analysis and histological characterization. A.B. performed the bioinformatic analysis. A.S., D.B. and T.V. performed scRNA-seq and analysis. M.F., A.V.K and J.-C.M. helped with breast cancer experiments. T.R. performed CTC experiments. S.B. helped with ATAC-seq. I.S., N.D'H., S.D.C., E.M., C.Bal., Y.S., A.L. and F.S. provided human samples. C.D. performed FACS sorting. V.M., S.L., F.D.C. and S.S. performed immunostaining, blocking antibody injections and follow-up of the mice. All authors read and approved the final manuscript.

METHODS

Compliance with ethical regulations. Mouse colonies were maintained in a certified animal facility in accordance with the European guidelines. All the experiments were approved by the corresponding ethical committee (Commission d'Etique et du Bien Etre Animal CEBEA, Faculty of Medicina, Universite Libre de Bruxelles). CEBEA follows the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS No.123). The mice were checked every day and were euthanized when the tumour reach the end-point size (1 cm in diameter or 1 cm^3 in volume) or if the tumour was ulcerated (independent of its size), if the mouse lost > 20% of its initial weight or showed any other sign of distress (based on general health status and spontaneous activity). None of the experiments performed in this study surpassed the size limit of the tumours. All the experiments complied strictly with the protocols approved by ethical committee. Patient-derived xenografts (PDX) used in this study to illustrate EMT transition states in human cancers are part of the PDX project, which has been approved by the ethical committee of the Hospital Erasme (Universite Libre de Bruxelles) and by ethical committees in all hospitals involved in patient recruitment. Informed consent was obtained from all patients.

Mouse strains. *Rosa26-YFP* mice[29], *Lgr5CreER* mice[30], *KRas^LSL-G12D* mice[31] and *p53^{fl/fl}* mice[32] were imported from the NCI mouse repository and Jackson Laboratories. *K8CreER* mice have previously been described[33]. *Pik3ca^{H1047R}* knock-in mice, in which wild-type exon 20 is replaced by H1047 mutant exon 20 upon Cre recombination, have previously been described[14]. W. Declercq (Ghent University, Belgium) generated the *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-ΔNp63-IRES-GFP* mice. *MMTV-PyMT* female mice were a gift from G. Berx (Ghent University) and M. Mazzone (VIB, University of Leuven). *NOD/SCID/Il2Rγ* null mice were purchased from Charles River. All mice used in this study were composed of males and females with mixed genetic background. No randomization and no blinding were performed in this study.

KRas^{LSL-G12D}p53^{fl/fl}-induced skin tumours. Tamoxifen was diluted at 25 mg/ml in sunflower seed oil (Sigma). Four daily intraperitoneal (IP) injections of 2.5 mg tamoxifen were administered at P28 as previously described[34] to *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-YFP^{+/+}* and *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-ΔNp63-IRES-GFP* mice. Tumour appearance and size were detected by daily observation and palpation. Mice were euthanized when tumour size reached 1 cm^3 or when mice presented signs of distress. Skin tumours were measured using a precision calliper, enabling us to identify 0.1-mm changes in size. Tumour volumes were measured on the first day of appearance of the tumour and then every week until the death of the animal with the formula $V = \pi[d^2D]/6$, in which d is the minor tumour axis and D is the major tumour axis.

K8CreER/Pic3ca^{H1047R}/p53^{fl/fl} and MMTV-PyMT breast tumours. Eight-week-old *K8CreER/Pik3ca^{H1047R}/p53^{fl/fl}* mice were induced with 15 mg tamoxifen (Sigma; diluted in sunflower seed oil) by IP injection (3 injections of 5 mg every 3 days). Hemizygous *MMTV-PyMT* females develop mammary tumours spontaneously. Tumour appearance and size were detected by daily observation and palpation. Mice were euthanized when a palpable mass of 1 cm^3 was detected.

Sample size, data exclusions, randomization and blinding. The sample size was chosen based on previous experience in the laboratory, for each experiment to yield high power to detect specific effects. No statistical methods were used to predetermine sample size.

For the metastasis analysis, only series of mice in which all subpopulations of TCs could be analysed were included. Those cases for which at least one grafted mouse died before the analysis were excluded. No other data were excluded from analysis.

For in vivo studies on primary mouse models, animals were chosen according to genotype, requiring four correct alleles (*Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-YFP^{+/+}* or *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-ΔNp63-IRES-GFP*). All animals were induced with tamoxifen at a similar age (28–35 days after birth), and the mice developed tumours in 2–3 months, thus minimizing the difference in age of the animals used. Sex-specific differences were minimized by including similar numbers of male and female animals if possible (for skin tumours). For breast cancer models, only female mice were used. Each experiment contained animals from at least three different litters. In experiments with anti-TGFβ, anti-Csf1r and anti-Ccl2 antibodies, animals from different litters were allocated randomly to the experimental and control groups. In the subcutaneous and intravenous grafting experiments we used *NOD/SCID/Il2Rγ* mice of similar ages and both female and male. For the circulating tumour cell experiments we screened all *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-YFP^{+/+}* mice and included in the analysis all animals in which we could detect CTCs.

Investigators were not blinded to mouse genotypes during experiments (for most of the experiments only one genotype was used: *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-YFP^{+/+}*). Researchers were not blinded when performing imaging and quantification. Two different researchers performed anti-Tgfβ, anti-Csf1r and anti-Ccl2 experiments; the researchers who analysed the FACS profiles of the tumours and histology were blinded during quantification and analyses.

Immunostaining. All stainings were performed on frozen sections. Tumour tissues and lungs were pre-fixed in 4% paraformaldehyde for 2 h at room temperature, rinsed in PBS, incubated overnight in 30% sucrose at 4 °C and embedded in OCT (Tissue Tek) for cryopreservation. Tissues were cut into 4–5-μm sections using a CM3050S (Leica Microsystems GmbH) cryostat and rinsed with PBS three times (5 min). Non-specific antibody binding was blocked with 5% horse serum, 1% BSA and 0.2% Triton X-100 for 1 h at room temperature. Primary antibodies were incubated overnight at 4 °C in blocking buffer. Sections were rinsed in PBS three times (5 min) and incubated with secondary antibodies diluted in blocking buffer at 1:400 for 1 h at room temperature. Nuclei were stained with Hoechst (4 mM) and slides were mounted using SafeMount (Labonord).

Antibodies for immunostaining. The following primary antibodies were used: anti-GFP (goat polyclonal, Abcam Cat#ab6673), anti-K14 (chicken polyclonal, Thermo Fisher Scientific Cat#MA5-11599), anti-vimentin (rabbit, clone ERP3776 Abcam Cat#ab92547), anti-CD61 (Armenian hamster, clone 2C9.G2, BD Bioscience Cat#553345), anti-CD31 (rat, clone MEC 13.3, BD Bioscience Cat#557355, rabbit polyclonal, Abcam Cat#ab28364), anti-CD45 (rat, clone 30-F11, BD Bioscience Cat#553079), anti-LYVE1 (rabbit polyclonal, Abcam cat#ab14917), anti-CD68 (rabbit polyclonal, Abcam Cat#ab125212), anti-S100a4/anti-Fsp1 (rabbit polyclonal, Thermo/Neomarkers, Cat# RB-9411-PO) and PYMT (rat monoclonal, Santa Cruz Biotechnology, Cat# sc-5381).

The following secondary antibodies were used: anti-rabbit, anti-rat, anti-goat, anti-chicken and anti-Armenian hamster conjugated to rhodamine Red-X (Jackson ImmunoResearch), Cy5 (Jakson ImmunoResearch) or Alexa Fluor-A488 (Molecular Probes).

Image acquisition. Image acquisition was performed on a Zeiss Axio Imager M1 (Thornwood) fluorescence microscope with a Zeiss Axiocam MR3 camera using Axiovision release 4.6 software. Brightness, contrast and picture size were adjusted using Adobe Photoshop CS6.

Screening of surface marker expression using BD Lyoplate. Skin tumours from *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-YFP^{+/+}* mice were dissected, minced and digested in collagenase type I (Sigma) at 3.5 mg/ml for 1 h at 37 °C on a rocking plate protected from light. Collagenase activity was blocked by the addition of EDTA (5 mM) and then the cells were rinsed in PBS supplemented with 10% FBS and the cell suspensions were filtered through a 70-μm cell strainer (BD Bioscience). Tumour cells were first incubated with APC-Cy7 conjugated anti-Epcam antibody (rat, clone G8.8, BD Cat#624073, dilution 1:50) for 30 min at 4 °C protected from light; next, a BD lyoplate screening panel for mouse cells (BD, Cat#562208), containing 176 anti-mouse monoclonal antibodies and 18 isotype controls, was applied. Living single tumour cells were selected by forward and side scatter, doublet discrimination and DAPI dye exclusion. The expression of the markers was analysed in YFP$^+$Epcam$^+$ and YFP$^+$Epcam$^-$ tumour cells (negative, homogeneously expressed or heterogeneously expressed).

Definition and FACS isolation of tumour cell subpopulations. Skin tumours from *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-YFP^{+/+}* mice and *Lgr5CreER/Kras^LSL-G12D/p53^{fl/fl}/Rosa26-ΔNp63-IRES-GFP* mice and secondary tumours arising from subcutaneously grafted tumour cells to *NOD/SCID/Il2Rγ*-null mice were dissected, minced and digested as described above. Breast tumours were dissected, minced and digested in 300 U ml^{-1} collagenase (Sigma) plus 300 μg ml^{-1} hyaluronidase (Sigma) for 2 h at 37 °C on a rocking plate protected from light. Brilliant violet stain buffer (BD Bioscience) was added (50 μl per sample) and the cells were incubated with PE-conjugated anti-CD51 (rat clone RMV-7, Biolegend Cat#104106, dilution 1:50), BV421-conjugated anti-CD61 (Armenian hamster, clone 2C9.G2, BD Bioscience Cat#553345, dilution 1:50), biotin-conjugated anti-CD106 (rat, clone 429 (MVCAM.A), BD Bioscience Cat#553331, dilution 1:50), BV711-conjugated anti-Epcam (rat clone G8.8, BD Bioscience Cat#563134, dilution 1:100), PerCPCy5.5-conjugated anti-CD45 (rat, clone 30-F11, BD Bioscience Cat#550994, dilution 1:100) and PerCPCy5.5-conjugated anti-CD31 (rat, clone MEC 13.3, BD Bioscience Cat#562861, dilution 1:100) antibodies for 30 min at 4 °C protected from light. For YFP$^-$ MMTV-PyMT tumours, BV605-conjugated anti-CD24 (rat, clone M1/69, Biolegend Cat#101827) and FITC-conjugated anti-CD29 (Armenian hamster, clone HMb1.1, Biolegend Cat#102206) antibodies were used to define tumour cells. Cells were washed with PBS supplemented with 2% FBS and incubated with streptavidin-BV786 (BD Bioscience Cat#563858, dilution 1:400) for 30 min at 4 °C protected from light. Living single tumour cells were selected by forward and side scatter, doublet discrimination and 7AAD dye exclusion. Tumour cells were selected by YFP expression and the exclusion of CD45 and CD31 (Lin$^-$). Different tumour cell subpopulations were defined in Epcam$^-$ tumour cells by a combination of CD61, CD51 and CD106 markers.

To screen for the heterogeneously expressed markers in different subpopulations, one of the following anti-mouse antibodies was used at dilution 1:50 in combination with staining for CD31, CD45, CDD106, CD51, CD61 and Epcam

as described above: AF647-conjugated anti-CD24 (clone M1/69 AF647 BD Cat#101817), AF647-conjugated anti-CD34 (clone RAM34 BD Cat#560233), APC-conjugated anti-CD43 (clone S11, Biolegend Cat#143207), APC-conjugated anti-CD44 (clone IM7, Biolegend Cat#103011), APC-conjugated anti-CD49b (clone DX5, Biolegend Cat#108909), APC-conjugated anti-CD54 (clone 3e2, BD Cat#561605), AF647-conjugated anti-CD71 (clone C2, BD Cat#563504), AF647-conjugated anti-CD104 (clone 346-11 A, Biolegend Cat#123607), APC-conjugated anti-CD140a (clone APA5 Biolegend Cat#135907), APC-conjugated anti-CD200 (clone OX-90, Biolegend Cat#123809), APC-conjugated anti-GITR/CD357 (clone DTA-1, Biolegend Cat#126311), APC-conjugated anti-podoplanin (clone 8.1.1, Biolegend Cat#127409) or APC-conjugated anti-Sca1 (clone E13-161.7, Biolegend Cat#122511). PE-conjugated anti-CD147 (clone RL73, BD Cat#562676) and anti-4-1BBl/CD137L (clone TSK-1, Biolegend Cat#107105) were used in combination with CD31, CD45, CD106, CD61 and Epcam, while CD51 was omitted (owing to the overlap of the fluorochromes).

FACS and analysis were performed using FACSAria and LSRFortessa, using FACSDiva software (BD Bioscience). Sorted cells were collected in culture medium (for in vivo transplantation experiments), lysis buffer (for RNA extraction) or PBS supplemented with 3% FBS (for ATAC).

Tumour transplantation assays. The different FACS-isolated tumour cell subpopulations (YFP$^+$Epcam$^+$, YFP$^+$Epcam$^-$CD106$^-$CD51$^-$CD61$^-$, YFP$^+$Epcam$^-$CD106$^+$CD51$^-$CD61$^-$, YFP$^+$Epcam$^-$CD106$^-$CD51$^+$CD61$^-$, YFP$^+$Epcam$^-$CD106$^+$CD51$^+$CD61$^-$, YFP$^+$Epcam$^-$CD106$^-$CD51$^+$CD61$^+$ and YFP$^+$Epcam$^-$CD106$^+$CD51$^+$CD61$^+$ from *Lgr5CreER/Kras$^{LSL-G12D}$/p53$^{fl/fl}$/Rosa26-YFP$^{+/+}$* mice) were collected in medium at 4 °C. Cells at different dilutions (1,000, 100 and 10 cells) were resuspended in 50 μl matrigel (E1270, 970 mg/ml; Sigma) and injected subcutaneously into *NOD/SCID/Il2rγ*-null mice. Secondary tumours were detected by palpation every week and their size monitored until tumours reached 1 cm^3 or when mice presented signs of distress. The mice were killed and the number of secondary tumours was quantified.

In vivo BrdU labelling of tumour cells. To evaluate the proliferative ability of different tumour cell subpopulations BrdU was injected IP at 50 mg/kg into *Lgr5CreER/Kras$^{LSL-G12D}$/p53$^{fl/fl}$/Rosa26-YFP$^{+/+}$* tumour-bearing mice at time points 0 and 12 h and the mice was killed at 24 h. The drinking water was supplemented with 1 mg/ml Brdu and 5% sucrose. Skin tumours were dissected, minced and digested as described above. APC BrdU Flow Kit (BD Cat#552598) was used according to the manufacturer's recommendations after surface marker labelling for CD31–PERCPCy5.5, CD45–PERCPCy5.5, CD106–BV786, CD51–PE, CD61–BV421 and Epcam–711). The percentage of BrdU-positive cells was analysed in each subpopulation (gated on living cells, CD45$^-$, CD31$^-$ negative and YFP$^+$).

In vitro invasion assay. The invasion ability of different tumour subpopulations was assessed using CultureCoat Cell Invasion Assay (RD Cat#3482-096-K) following the manufacturer's recommendations. In brief, 25,000 freshly sorted cells resuspended in medium with a low concentration of serum (0.5%) were placed in the top chamber of the cell invasion device, while medium with 20% serum was added to the lower chamber and left at 37 °C for 24 h. Then the medium was aspirated and the chambers were washed with wash buffer. Calcein was diluted in cell dissociation solution, 100 μl were added to the bottom chamber and the cell invasion device was incubated at 37 °C for 1 h. Then the top chamber was removed and the bottom plate was read at 485-nm excitation and 520-nm emission.

Metastasis assay. The different FACS-isolated tumour cell subpopulations were collected in medium at 4 °C. Cells were resuspended in 50 μl PBS and injected into the tail veins of *NOD/SCID/Il2Rγ*-null mice (500 cells per injection for skin SCCs, 2,000 cells for *K8CreER/Pik3ca/p53$^{fl/fl}$* breast tumours and 10,000 cells for *MMTV-PyMT* breast tumours). Mice were killed after 40 days and the lungs were analysed to detect the metastases. The number of metastases was quantified on 10 cryosections per lung (separated by 100 μm) based on YFP (skin SCCs and *K8CreER/Pik3ca/p53$^{fl/fl}$* metaplastic breast tumours) or PYMT (*MMTV-PyMT* breast tumours) expression and presented as number of metastases per lung. The metastases were characterized for K14 and vimentin expression by immunofluorescence.

Detection of circulating tumour cells. An average of 1 ml blood was extracted by cardiac puncture with 50 μl EDTA (0.5 M, pH 8, Ambion P/N: AM9260G, L/N: 1509002). Red blood cells were lysed with 0.2% NaCl (VWR chemicals) and brought to isotonic condition with 2.6% NaCl, 0.2% sucrose (MERCK KGaA). The blood was incubated with collagenase at 37 °C for 30 min to dissociate CTC clusters. FACS staining and analysis were performed as described for tumours. The results are presented as average percentage of different CTC subpopulations per mouse.

In vivo anti-TGFβ treatment. Five weeks after tamoxifen administration (before tumour appearance), *Lgr5CreER/Kras$^{LSL-G12D}$/p53$^{fl/fl}$/Rosa26-YFP$^{+/+}$* mice were injected IP with 500 μg neutralizing anti-TGFβ (clone 1D11.16.8, Bio X Cell Cat#BE0057) or control (clone MOPC-21, Bio X Cell Cat#BE0083) antibodies every 4 days during the first week, followed by 200 μg every 2 days until the tumour size reached 1 cm^3 or mice presented signs of distress.

In vivo macrophage depletion treatment. Five weeks after tamoxifen administration (before tumour appearance), *Lgr5CreER/Kras$^{LSL-G12D}$/p53$^{fl/fl}$/Rosa26-YFP$^{+/+}$* mice were injected IP with 300 μg neutralizing anti-Csf1r (clone AFS98, Bio X Cell Cat#BE0213) and 200 μg anti-Ccl2 (clone 2H5, Bio X Cell Cat# BE0185) antibodies every 4 days or with control antibodies (clone 2A3, Bio X Cell Cat#BE0089) until the tumour size reached 1 cm^3 or mice presented signs of distress.

RNA extraction and real-time PCR. RNA was extracted from FACS-isolated cells using RNeasy micro kit (QIAGEN) according to the manufacturer's recommendations. For real-time PCR, after nanodrop quantification, the first-strand cDNA was synthesized using Superscript II (Invitrogen) and random hexamers (Roche) in 50 μl final volume. Control of genomic contamination was measured for each sample by performing the same procedure with or without reverse transcriptase. Quantitative PCR assays were performed using 1 ng cDNA as template, SYBRGreen mix (Applied Bioscience) on a Light Cycler 96 (Roche) real-time PCR system. *HPRT* housekeeping gene was used for normalization. Primers were designed using a Pubmed tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) or Roche Universal Probe Library Assay Design centre (https://lifescience.roche.com/en_es/articles/Universal-ProbeLibrary-Assay-Design.html).

RNA sequencing. Before sequencing the quality of RNA was evaluated by Bioanalyzer 2100 (Agilent). Indexed cDNA libraries were obtained using Ovation Solo RNA-seq Systems (NuGen) following the manufacturer's recommendations. The multiplexed libraries (11 pM/18 pM) were loaded onto flow cells and sequences were produced using a HiSeq PE Cluster Kit v4 and TruSeq SBS Kit v3-HS (250 cycles) on a HiSeq 1500 (Illumina). Approximately 8 million paired-end reads of 75 bp per sample for each subpopulation sample and approximately 12 million paired-end reads of 125 bp per sample for samples overexpressing p63 were mapped against the mouse reference genome (Grcm38/mm10) using STAR software[35] to generate read alignments for each sample. Annotations for Grcm38.87 were obtained from ftp://ftp.ensembl.org/. After transcript assembly, gene level counts were obtained using HTseq and normalized to 20 million aligned reads. Average expression for each gene for the different tumour cell subpopulations was computed using three biological replicates and fold changes were calculated between the subpopulations. Genes for which all the mean expressions across the subpopulations were lower than 1 read per million mapped reads were considered not expressed and removed from further analysis. For some comparisons, we took a threshold of 10 reads per million. Genes having a fold change of expression greater than or equal to 2 were considered upregulated and those having a fold change of expression lower than or equal to 0.5 were considered downregulated.

Single-cell RNA sequencing. Single-cell transcriptomes were generated using a modified Smartseq-2 protocol. Lysis buffer (1 μl) in 384-well PCR plates for cell sorting was prepared with 0.4% v/v Triton-X lysis buffer, 2.5 mM dNTPs, 2.5 μM oligo-dT30-VN and ERCC controls at a final dilution of 1:400 million. Reverse transcription and PCR were performed according to the Smartseq-2 protocol with reduced volumes: 1 μl reverse transcription mix instead of 5.7 μl and 5 μl PCR Master mix instead of 12.5 μl. cDNA was amplified for 24 cycles and cleaned using HighPrep PCR beads (MagBio Genomics) at a 0.8 × ratio on a Hamilton STAR (Hamilton Germany GmbH) liquid handler, eluted in 30 μl buffer EB (Qiagen) and transferred to 384_PP acoustic plates (LabCyte). DNA quantification was performed with Picogreen assay (Thermo Fisher Scientific) and a subset of samples were selected for quality control using a Agilent 2100 BioAnalyser (Agilent Technologies) using a High Sensitivity DNA kit.

Library preparation continued from Smartseq-2. Next-generation sequencing library preparation was performed using the Illumina Nextera DNA library preparation kit with volumes reduced by one-tenth (Illumina). In brief, 100 pg cDNA in a volume of 500 nl was tagmented by adding 1.5 μl Tn5-buffer mix and incubating for 10 min at 55 °C. Tagmented samples were barcoded with Nextera indexes A–D and amplified with 11 cycles of PCR. After PCR, all samples were pooled and cleaned using HighPrep PCR beads at a 0.6 × ratio. Library pools were eluted in buffer EB and quality controlled using an Agilent 2100 BioAnalyser and High Sensitivity DNA kits before adjusting to a concentration of 4 nM. The diluted pools were quantified using the KAPA qPCR library quantification kit on a LightCycler 480 (Roche) before final dilution to 2 nM. Sequencing was performed on a HiSeq2500 on two lanes in high output mode.

Single-cell bioinformatics analysis. Sequencing reads were trimmed for adaptor sequences using cutadapt (version 1.13) and reads were aligned to the GRCm38 reference genome including ERCC sequences using STAR with default parameters (version 2.5.2b). The expression count matrix was generated using HTSeq (version 0.6.0) on GENCODE M12 transcript annotations and counts for each protein coding gene were collapsed. Quality control was performed using the scater R package (version 1.2.0). Cells that complied with one of the following conditions were excluded: had fewer than 10^5 counts, showed expression of fewer than 2,000 unique genes, had more than 20% counts belonging to ERCC sequences, or had more than 10% counts belonging to mitochondrial sequences. Genes for which fewer than 20 counts were observed across the complete dataset were excluded

from further analysis. Read counts were normalized using scran with default parameters (version 1.2.2). We assigned cell-cycle phase using cyclone from the scran R package and omitted all cells that were not in the G1 phase from further analysis. Clustering using the SC3 R package (version 1.3.18)[36] and PCA was performed using the prcomp function in R and plots were generated using the ggplot2 R package (version 2.2.1). We chose $k = 5$ for SC3 as this best represented the heterogeneity in our dataset and recapitulated the EMT states with the clusters representing transitional states. For cluster marker gene discovery we set the thresholds to all genes with an AUC higher than 0.8 and $P < 0.01$. Heat maps were generated using a modified version of the gplots R package (3.0.1). Pseudotime-ordering was performed using the monocle[37] package (version 3.2.6) by first performing dimensionality reduction using tSNE on a subset of the top 500 genes with the highest covariability with CD51, CD61 and CD106 marker state. This state was determined by simple rule-based logic where positive status was assumed if more than one read for CD51, CD61 or CD106 was observed. We then performed clustering ($k = 5$) and identified the top 2,000 genes with the highest cluster specificity. These genes were used to compute the pseudotime-ordering and the cell trajectory. We then looked for genes found to be differentially expressed across the branch point in the trajectory using the BEAM approach, as implemented in monocle.

ATAC sequencing. For ATAC-seq, 100,000 sorted cells from different tumour subpopulations were collected in 1 ml PBS supplemented with 3% FBS at 4 °C. Cells were centrifuged and cell pellets were resuspended in 100 μl lysis buffer (TrisHCl 10 mM, NaCl 10 mM, MgCl2 3 mM, Igepal 0.1%) and centrifuged at 500g for 25 min at 4 °C. Supernatant was carefully discarded and nuclei were resuspended in 50 μl reaction buffer (Tn5 transposase 2.5 μl, TD buffer 22.5 μl, from Nextera DNA sample preparation kit, Illumina, and 25 μl H$_2$0). The reaction was performed at 37 °C for 30 min and was stopped by addition of 5 μl clean up buffer (NaCl 900 mM, EDTA 300 mM). DNA was purified using the MiniElute purification kit (QIAGEN) following the manufacturer's protocol. DNA libraries were PCR amplified (Nextera DNA Sample Preparation Kit, Illumina), and size selected from 200 to 800 bp (BluePippin, Sage Sciences), following the manufacturer's recommendations.

ATAC-seq analysis. Adaptor sequences were removed with TrimmomaticPE using options 'ILLUMINACLIP:adaptor.file:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36'. ATAC–seq paired-end reads were then aligned to mouse genome Grcm38 using Bowtie2 (version 2.2.6)[38] using options '-X 2000–fr–very-sensitive–no-discordant–no-unal–no-mixed –non-deterministic'. Mitochondrial reads, reads from unmapped or random contigs and reads with a mapping quality < 20 were removed using samtools[39]. Duplicate reads were removed by Picard tools (http://broadinstitute.github.io/picard/).

Peak calling was performed on each individual sample by macs2 (version 2.1.0.20151222)[40] using options '-f BAMPE -g mm -q 0.05–nomodel–call-summits -B –SPMR'. Peaks from the different subpopulations were merged for downstream analysis.

Reads counts of each merged peak for each individual sample were calculated by HTSeq-count[41] using options '-s no -m intersection-nonempty'. These counts were normalized for one million mapped reads in merged peaks and fold-change was calculated between the subpopulations. Peaks were associated to genes with GREAT software[42] with the following parameters: 5.0 kb in proximal upstream, 1.0 kb in proximal downstream and 100.0 kb in distal. For most of the analysis, only peaks annotated to one gene were kept.

Differential peaks are defined as peaks having at least a twofold change between the two subpopulations and being called peak in the subpopulation where they are higher.

De novo motif search was performed using findMotifsGenome.pl program in HOMER software[43] using parameters 'size −250,250 -S 15 -len 6,8,10,12,16.' Specific motif research was performed using annotatePeaks.pl program in HOMER software using parameters '-size 500'.

Quantification and statistical analysis. Two-tailed Student's t-test, two-tailed Mann–Whitney U-test and Kaplan–Maier survival analysis were performed using GraphPad Prism version 7.00 for Mac (GraphPad Software).

The TPC frequency was computed using extreme limiting dilution analysis (ELDA) online software as previously described (http://bioinf.wehi.edu.au/software/elda/)[44]. The statistical P value was obtained using a χ^2 test. The statistical P values for the number of metastases and the proportion of the different subpopulations in skin and metaplastic breast tumours were calculated using a t-test.

Other statistical methods: dendrogram for clustering of subpopulations (RNA-seq and ATAC-seq data) was drawn on the number of reads in the 500 annotated merged peaks with highest variance across the samples using canberra distance and complete clustering method with heatmap.2 function in R software (Foundation for Statistical Computing; http://www.bioconductor.org/).

To estimate whether the proportion of peaks annotated to genes having a smad2 binding site for all the different comparisons of EMT and MET was higher than in the whole genome, we randomly generated five sets of 10,000 peaks of 500-bp in the whole mouse genome with RSAT software[45]. Smad2 binding sites were searched with HOMER exactly as in peaks for EMT leading to 2,832, 2,788, 2,782, 2,764 and 2,797 peaks among the 10,000 having at least one potential binding site for Smad2. We took 2,793 peaks as the average. The prop.test function of R software was used with the 'alternative = greater' parameter to test whether the proportion was higher.

All the statistical analyses are based on biological replicates (n indicated in the text, figures or figure legends). No technical replicates were used to calculate statistics.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. Data associated with this study have been deposited in the NCBI Gene Expression Omnibus under accession numbers GSE110587 (RNA-seq and ATAC-seq) and GSE110357 (scRNA-seq).

29. Srinivas, S. et al. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4 (2001).
30. Barker, N. et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007).
31. Tuveson, D. A. et al. Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. *Cancer Cell* **5**, 375–387 (2004).
32. Jonkers, J. et al. Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat. Genet.* **29**, 418–425 (2001).
33. Van Keymeulen, A. et al. Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189–193 (2011).
34. Lapouge, G. et al. Skin squamous cell carcinoma propagating cells increase with tumour progression and invasiveness. *EMBO J.* **31**, 4563–4575 (2012).
35. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
36. Kiselev, V. Y. et al. SC3: consensus clustering of single-cell RNA-seq data. *Nat. Methods* **14**, 483–486 (2017).
37. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
38. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
39. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
40. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
41. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
42. McLean, C. Y. et al. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
43. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
44. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cells and other assays. *J. Immunol. Methods* 15, (70–78 (2009).
45. Medina-Rivera, A. et al. RSAT 2015: regulatory sequence analysis tools. *Nucleic Acids Res.* **43**, W50–W56 (2015).
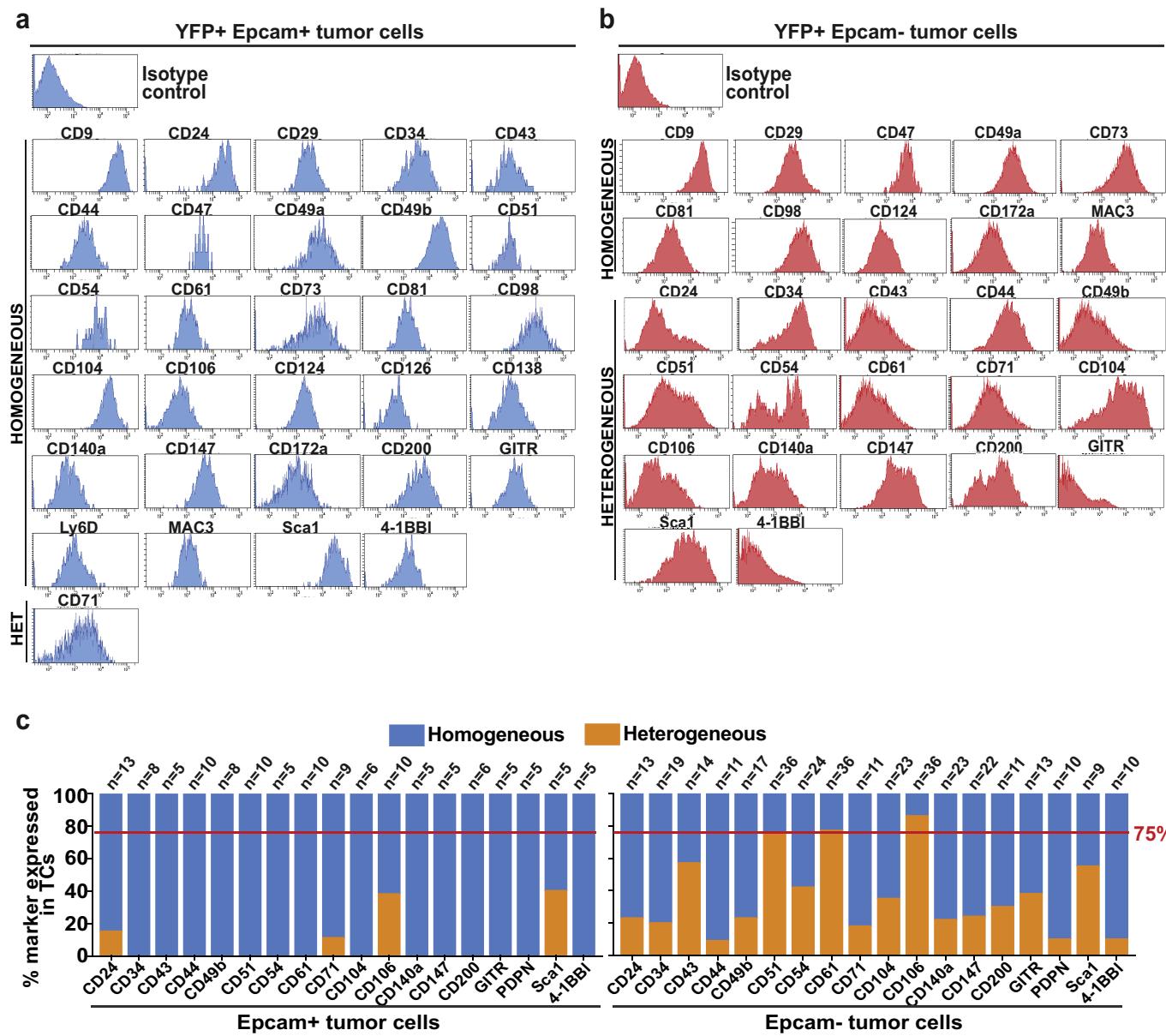
Extended Data Fig. 1 | Screening of cell surface markers homogeneously and heterogeneously expressed in Epcam⁺ and Epcam⁻ TCs in skin SCCs. a, **b**, FACS profile of the isotype control and homogeneously or heterogeneously expressed markers in Epcam⁺ (**a**) and Epcam⁻ (**b**) TCs using BD lyoplate panel. **c**, Screening of all markers found to be heterogeneously expressed in at least one tumour in the initial BD lyoplate screen on a much larger cohort of tumours. Epcam⁻ TCs were more heterogeneous than Epcam⁺ TCs and CD51, CD61 and CD106 were heterogeneously expressed in at least 75% of the tumours analysed and therefore were selected and used in combination for further experiments.
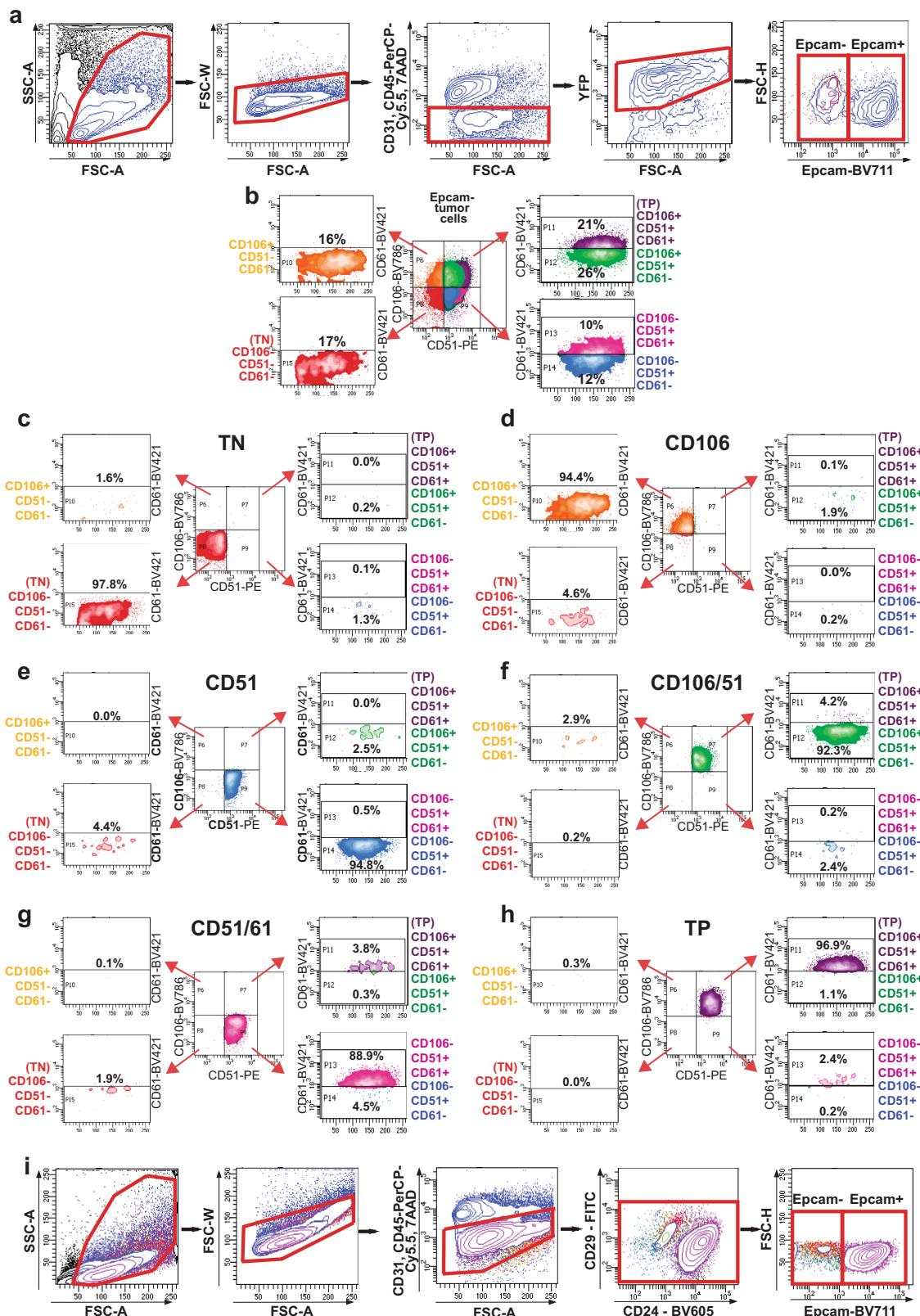
Extended Data Fig. 2 | Gating strategy for FACS analysis and cell sorting of the different tumour subpopulations. a, FACS plots showing the gating strategy used to FACS-isolate Epcam[+] and Epcam[−] cells from YFP[+] tumours (*Lgr5CreER/Kras*^{G12D}*/p53*^{KO} and *K8CreER/ Pik3ca/p53*^{KO}). **b**, FACS plots showing the six different subpopulations of Epcam[−] TCs: Epcam[−]CD106[−]CD51[−]CD61[−] (triple-negative), Epcam[−]CD106[+]CD51[−]CD61[−], Epcam[−]CD106[−]CD51[+]CD61[−],

Epcam[−]CD106[+]CD51[+]CD61[−], Epcam[−]CD106[−]CD51[+]CD61[+] and Epcam[−]CD106[+]CD51[+]CD61[+] (triple-positive) populations. **c–h**, Post-sort purity plots for triple-negative (**c**), CD106[+] (**d**), CD51[+] (**e**), CD106[+]CD51[+] (**f**), CD51[+]CD61[+] (**g**) and triple-positive (**h**) subpopulations (*n* = 1). **i**, FACS plots showing the gating strategy used to FACS-isolate Epcam[+] and Epcam[−] cells from *MMTV-PyMT* mammary tumours.
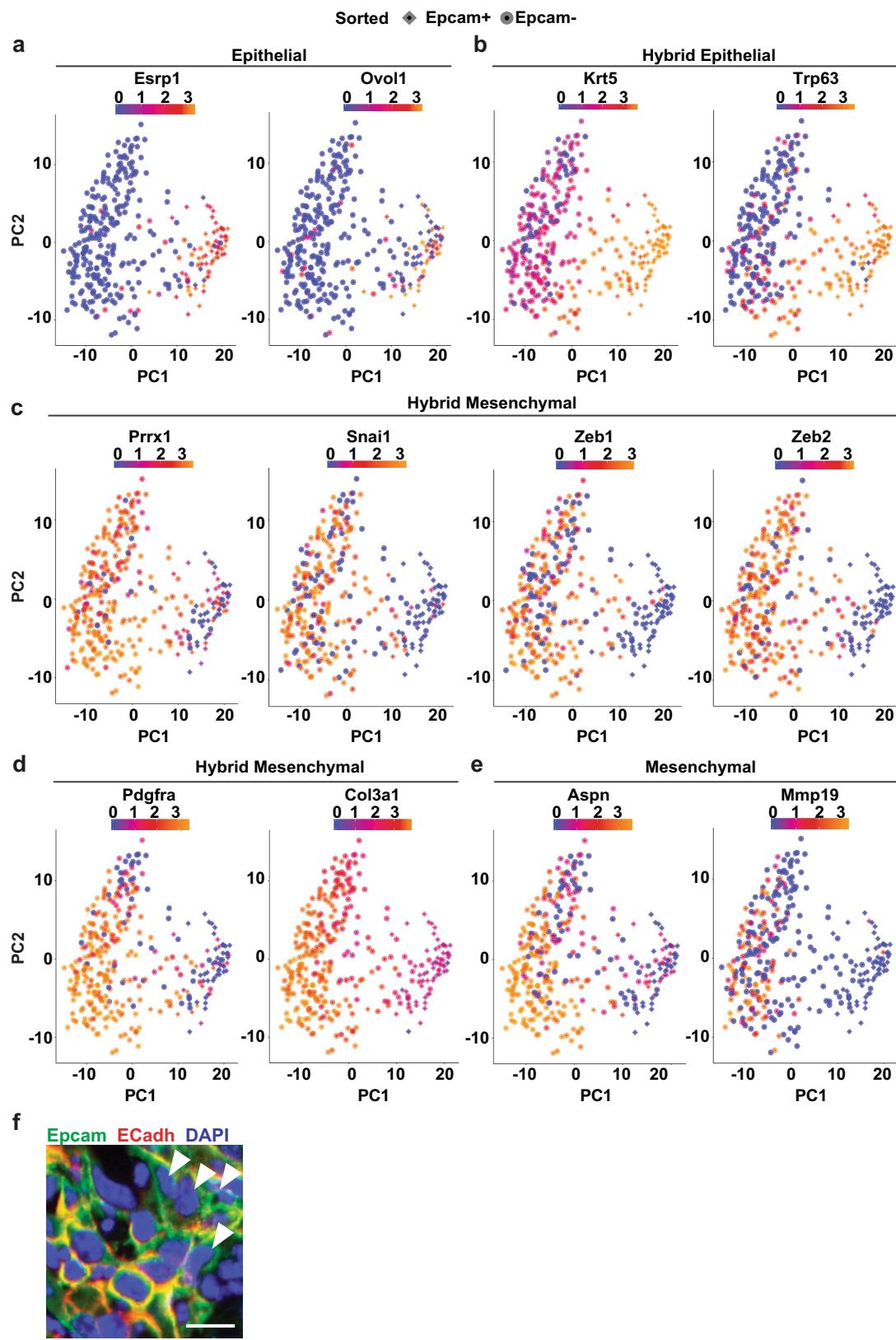
Extended Data Fig. 3 | Epithelial, hybrid and mesenchymal marker expression by scRNA-seq. a–e, Dimensionality reduction of scRNA-seq data using PCA colouring for pure epithelial genes *Esrp1* and *Ovol1* (**a**), hybrid epithelial genes *Krt5* and *Trp63* (**b**), hybrid mesenchymal genes *Prrx1, Snai1, Zeb1, Zeb2, Pdgfra* and *Col3a1* (**c, d**) and pure mesenchymal genes *Aspn* and *Mmp19* (**e**). Every dot represents a single cell and its shape represents the FACS sorting criteria (Epcam$^+$ or Epcam$^-$). The colour scale represents the normalized expression of the respective genes. Pure epithelial genes are expressed only in Epcam$^+$ TCs, hybrid epithelial genes in both Epcam$^+$ and hybrid Epcam$^-$ TCs, hybrid mesenchymal genes in hybrid Epcam$^-$ TCs and highly mesenchymal Epcam$^-$ TCs, and pure mesenchymal genes only in highly mesenchymal Epcam$^-$ TCs. **a–e,** $n = 66$ Epcam$^+$ and $n = 277$ Epcam$^-$ cells from one tumour. **f,** Immunostaining showing the co-expression of Epcam and E-cadherin in the majority of epithelial TCs. Rare Epcam + cells are E-cadherin$^-$ (arrowheads). Scale bars, 50 μm.
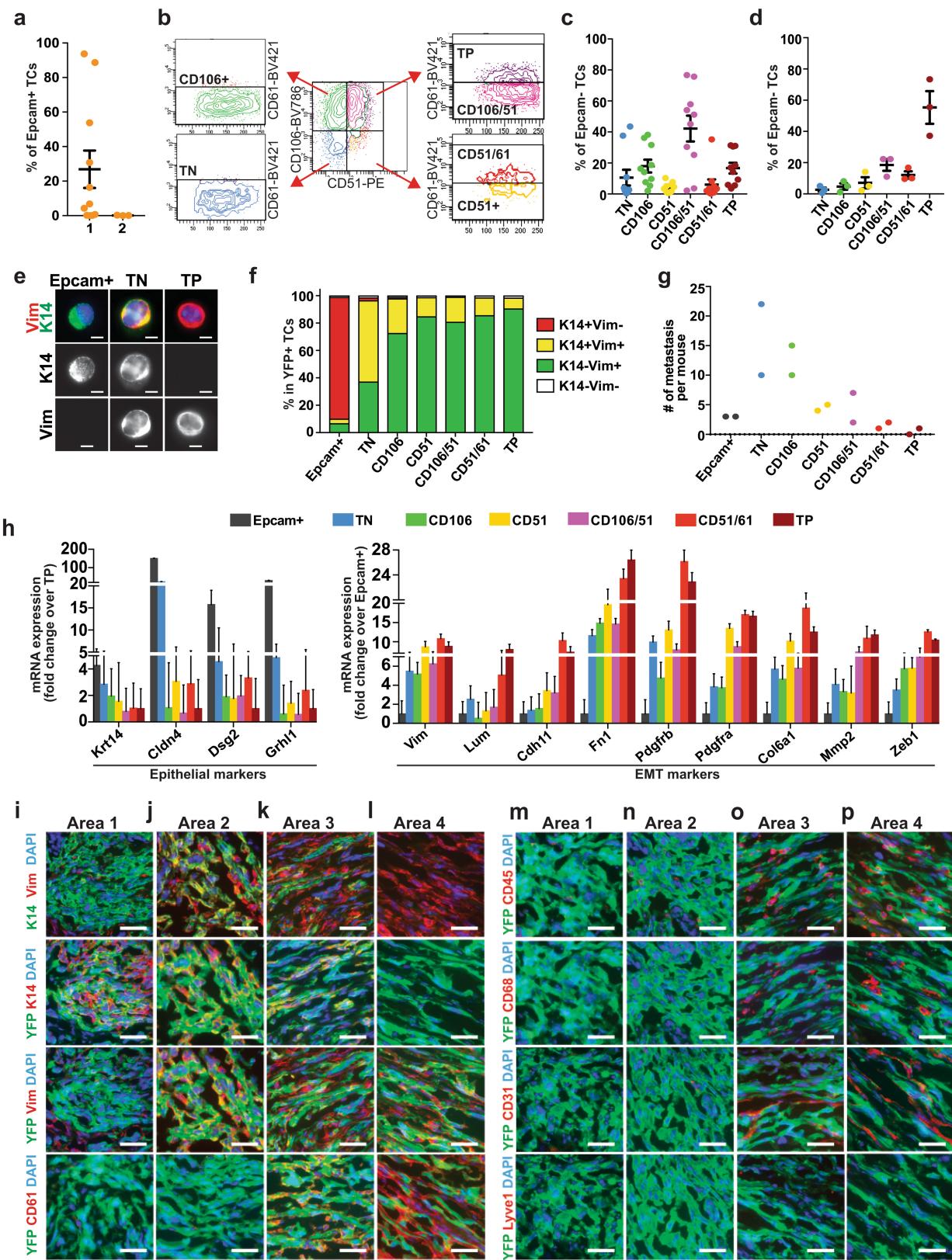
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | EMT transition states in *Pik3ca^{H1047R}/p53^{KO}* metaplastic mammary tumours. a, Percentage of Epcam$^+$ cells in the two groups of tumours that differ by the frequency of the different CD106, CD51, CD61 subpopulations in metaplastic-like breast tumours resulting from activation of the oncogenic mutation *Pik3ca^{H1047R}* and *p53* deletion in luminal cells of the mammary gland ($n = 11$ biological replicates, mean ± s.e.m.). **b**, FACS profile showing the six different CD106, CD51, CD61 subpopulations in the most frequent group 1 mixed tumours containing Epcam$^+$ and Epcam$^-$ TCs. **c, d**, Distribution of the different tumour subpopulations in group 1 (**c**; mixed tumours with Epcam$^+$ and Epcam$^-$ TCs) and in group 2 (**d**; mesenchymal Epcam$^-$ TCs) ($n = 10$ biological replicates, mean ± s.e.m.). **e**, Immunostaining of K14 and vimentin in cystospin of FACS-isolated YFP$^+$Epcam$^+$, YFP$^+$Epcam$^-$ triple-negative and YFP$^+$Epcam$^-$ triple-positive populations. Scale bars, 20 μm; $n = 3$ biological replicates. **f**, Proportions of cells expressing K14 and vimentin based on cytospin counts; $n = 3$ biological replicates. **g**, Number of lung metastases arising from the injection of 2,000 YFP$^+$ TCs from different EMT subpopulations. Dots represent the number of metastases per condition and per mouse ($n = 2$). **h**, qRT–PCR showing the downregulation of epithelial markers and the upregulation of EMT markers in the different Epcam$^-$ tumour subpopulations. Histogram represents mean (+ s.e.m.) fold-change in gene expression compared to Epcam$^+$ tumour cells for mesenchymal genes and compared to triple-positive tumour cells for epithelial genes ($n = 7$ Epcam$^+$, $n = 3$ triple-negative, $n = 5$ CD106$^+$, $n = 3$ CD51$^+$, $n = 6$ CD106$^+$CD51$^+$, $n = 3$ CD51$^+$CD61$^+$, $n = 6$ triple-positive, all are biological replicates). **i–l**, Immunostaining showing the expression of K14 and vimentin, YFP and K14, YFP and vimentin, or YFP and CD61 in area 1 (K14$^+$Vim$^-$CD61$^-$; **i**), area 2 (K14$^+$Vim$^+$CD61$^-$; **j**), area 3 (K14$^+$Vim$^+$CD61$^+$; **k**) and area 4 (K14$^-$Vim$^+$CD61$^+$; **l**). Scale bars, 50 μm. **m–p**, Immunostaining showing the expression of YFP and CD45, CD68, CD31 or Lyve1 in area 1 (**m**), area 2 (**n**), area 3 (**o**) and area 4 (**p**), showing that immune infiltration and the density of vascular and lymphatic vessels are associated with late stages of EMT. Scale bars, 50 μm, $n = 3$ biological replicates. In hair follicle EMT-derived tumours, the chromatin and transcriptional landscape of the cancer cell of origin prime the tumour to undergo EMT during tumorigenesis; in *Pik3ca/p53^{KO}* mammary metaplastic tumours this is not always the case. Metaplastic EMT tumours are more frequent in luminal-derived tumours than in basal/myoepithelial cells, which under physiological conditions present much more EMT-like features do than luminal cells.
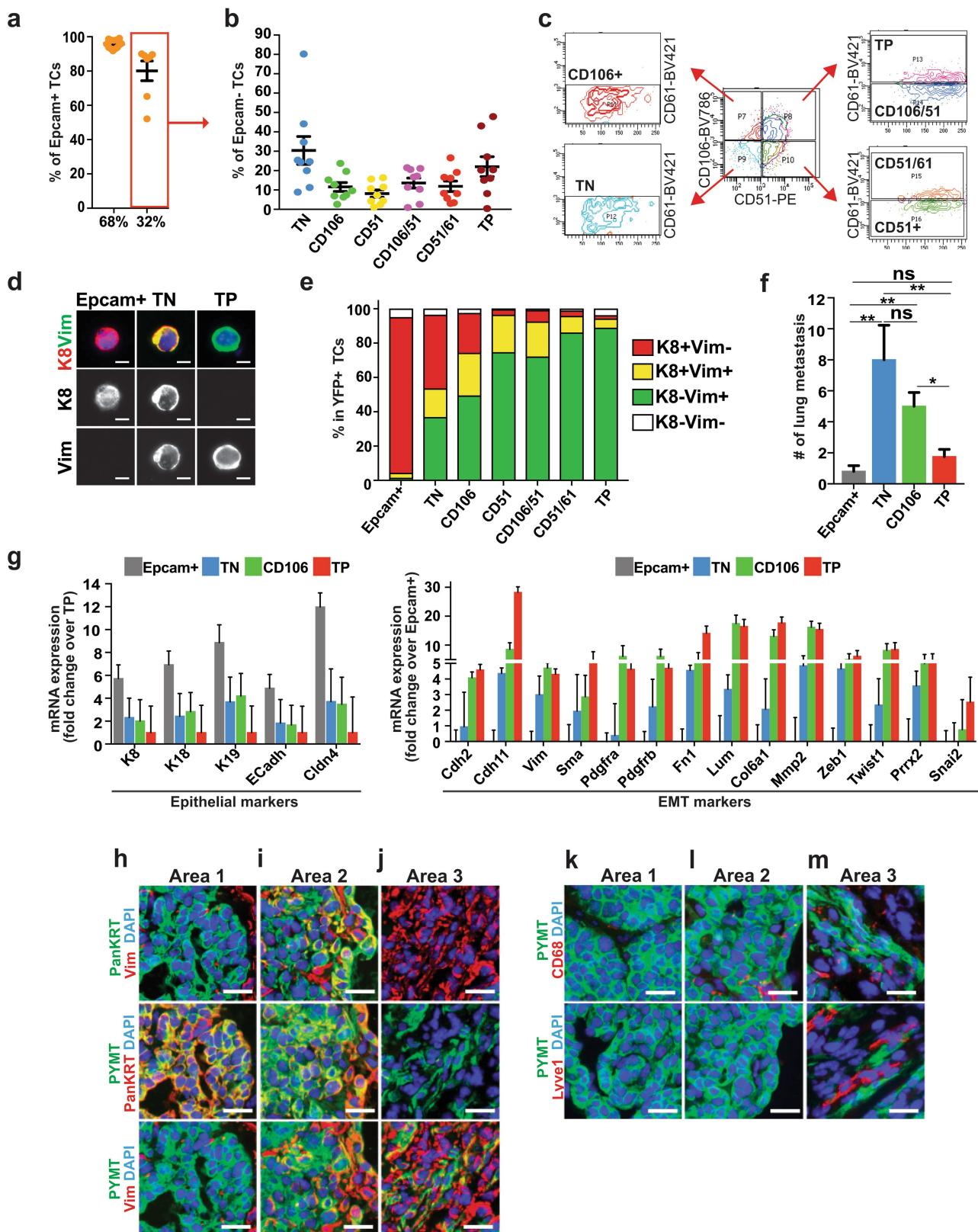
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | EMT transition states in *MMTV-PyMT* luminal-like mammary tumours. a, Percentage of Epcam$^+$ cells in two groups of tumours that differ by the frequency of the different CD106, CD51, CD61 subpopulations in *MMTV-PyMT* luminal-like mammary tumours ($n = 5$ mice, mean $\pm$ s.e.m.). **b**, Distribution of the different tumour subpopulations in group 2 (mixed tumours with Epcam$^+$ and Epcam$^-$ TCs) ($n = 9$ tumours, $n = 5$ mice, mean $\pm$ s.e.m.). **c**, FACS profile showing the six different CD106, CD51, CD61 subpopulations in the group containing mixed tumours with Epcam$^+$ and Epcam$^-$ TCs. **d**, Immunostaining of K8 and vimentin in cystospin of FACS-isolated YFP$^+$Epcam$^+$, YFP$^+$Epcam$^-$ triple-negative and YFP$^+$Epcam$^-$ triple-positive populations. Scale bars, 20 μm, $n = 2$ biological replicates. **e**, Proportion of cells expressing K8 and vimentin based on cytospin counts ($n = 2$ tumours). **f**, Number of lung metastases arising from the injection of 10,000 TCs from the different EMT subpopulations (Epcam$^+$, $n = 5$; triple-negative, $n = 5$; CD106$^+$, $n = 5$; triple-positive, $n = 4$). All are biological replicates, mean + s.e.m. Two-sided Mann–Whitney test:

Epcam$^+$ versus triple-negative, $P = 0.0079$; Epcam$^+$ versus CD106$^+$, $P = 0.0079$; triple-negative versus triple-positive, $P = 0.0238$; CD106$^+$ versus triple-positive, $P = 0.0476$; mean + s.e.m. **g**, qRT–PCR showing the downregulation of epithelial markers and the upregulation of EMT markers in the different Epcam$^-$ tumour subpopulations. Histogram represents mean $\pm$ s.e.m. fold-change in gene expression compared to Epcam$^+$ tumour cells for mesenchymal genes and compared to triple-positive tumour cells for epithelial genes (Epcam$^+$, $n = 5$; triple-negative, $n = 3$; CD106$^+$, $n = 5$; triple-positive, $n = 4$; all are biological replicates). **h–j**, Immunostaining of pancytokeratin (PanKRT) and vimentin, PyMT and pancytokeratin or PyMT and vimentin in area 1 (pancytokeratin$^+$Vim$^-$; **h**), area 2 (pancytokeratin$^+$Vim$^+$; **i**) and area 3 (pancytokeratin$^-$Vim$^+$; **j**). Scale bars, 50 μm. **k–m**, Immunostaining of PyMT and CD68 or Lyve1 in area 1 (**k**), area 2 (**l**) and area 3 (**m**), showing that immune infiltration and the density of lymphatic vessels are associated with late-stage EMT. Scale bars, 50 μm, $n = 2$ biological replicates.
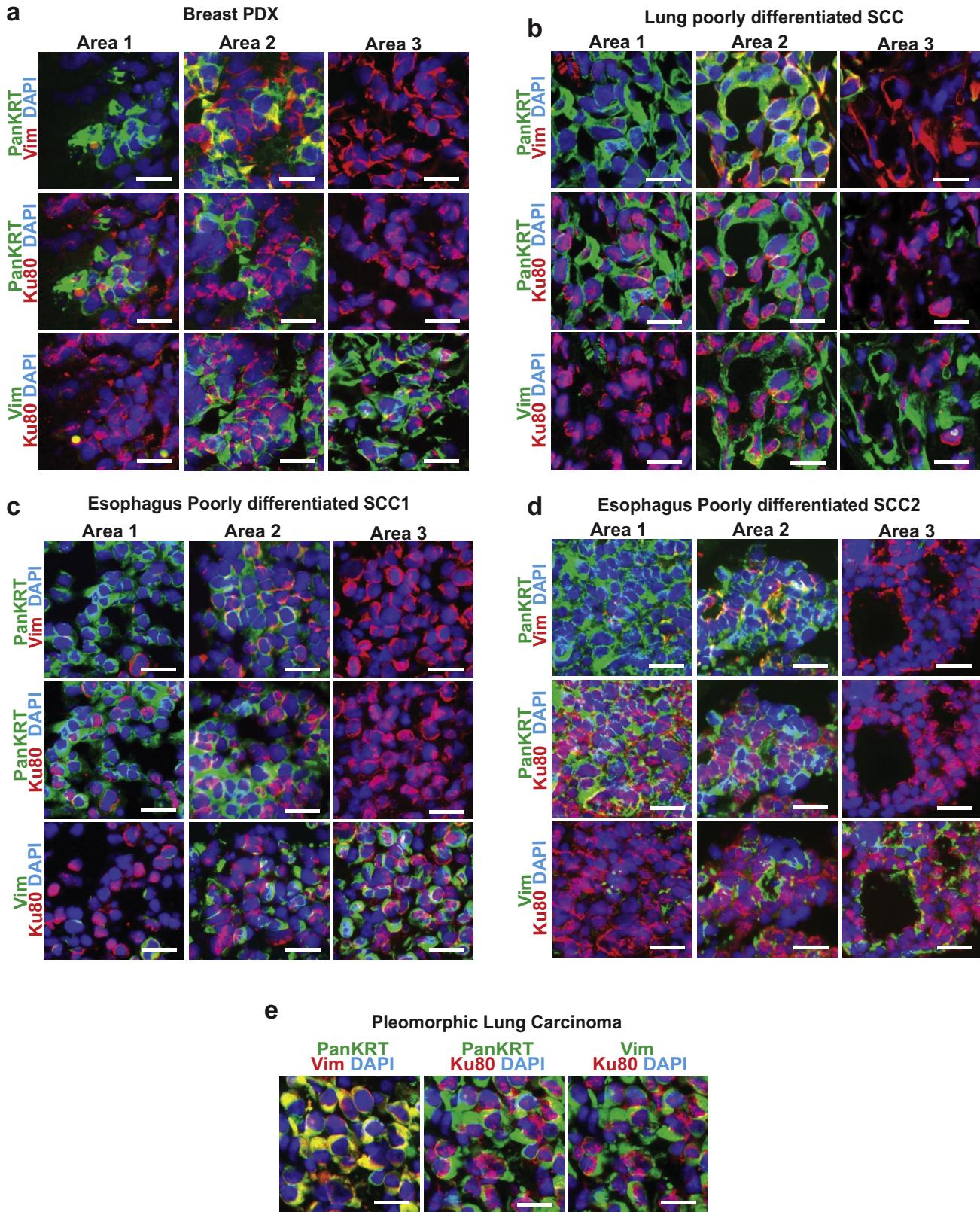
Extended Data Fig. 6 | EMT transition states in human patient-derived xenotransplantation (PDX). a–d, Immunostaining of pancytokeratin and vimentin, human-specific marker Ku80 and pancytokeratin or Ku80 and vimentin in poorly differentiated human breast (**a**), lung (**b**) or oesophagus (**c, d**) PDX. In these examples, three different areas can be identified: area 1 (Ku80[+]pancytokeratin[+]vimentin[−]), area 2 (Ku80[+]pancytokeratin[+]vimentin[+]) and area 3 (Ku80[+]pancytokeratin[−]vimentin[+]), showing that TCs with different degrees of EMT can be identified in a subset of poorly differentiated human cancers. Scale bars, 50 μm. **e,** PDX derived from aggressive and highly metastatic pleomorphic lung carcinoma, showing the hybrid EMT state as defined by co-expression of pancytokeratin and vimentin by the majority of TCs.
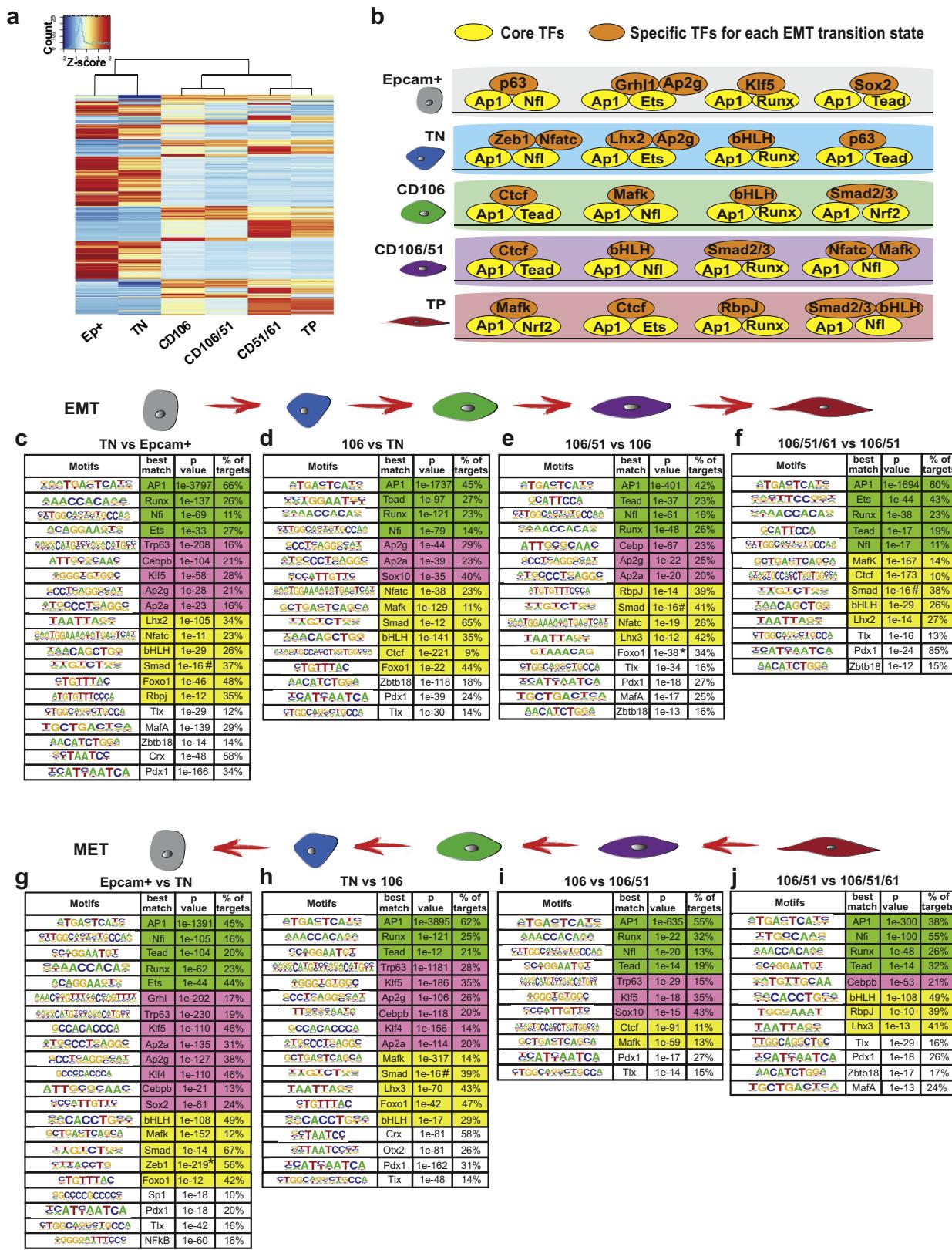
Extended Data Fig. 7 | Transcription factor binding motif enrichment in opened chromatin regions in the different EMT populations.
a, Heat map representing unsupervised clustering of RNA-seq data from the different tumour subpopulations, showing that Epcam+ epithelial TCs, Epcam− hybrid TCs and Epcam− mesenchymal TCs fall into different clusters ($n = 3$ biological replicates). **b**, Representation of chromatin remodelling and associated transcription factors (TFs) during EMT based on the transcription factor motifs that were statistically enriched in ATAC-seq peaks that differed between tumour subpopulations as determined by Homer analysis. **c–f**, Transcription factor motifs enriched

in the ATAC-seq peaks that were upregulated between the indicated subpopulations as determined by Homer analysis using known or de novo motif search (asterisk) or using JASPAR motif matrix (hash symbol). **g–j**, Transcription factor motifs enriched in the ATAC-seq peaks that were upregulated between the indicated tumour subpopulations as determined by Homer analysis using known or de novo motif search (asterisk). Green, core transcription factors; pink, epithelial transcription factors; yellow, mesenchymal transcription factors. **c–j**, Analysis based on ATAC-seq results from different subpopulations derived from one tumour.
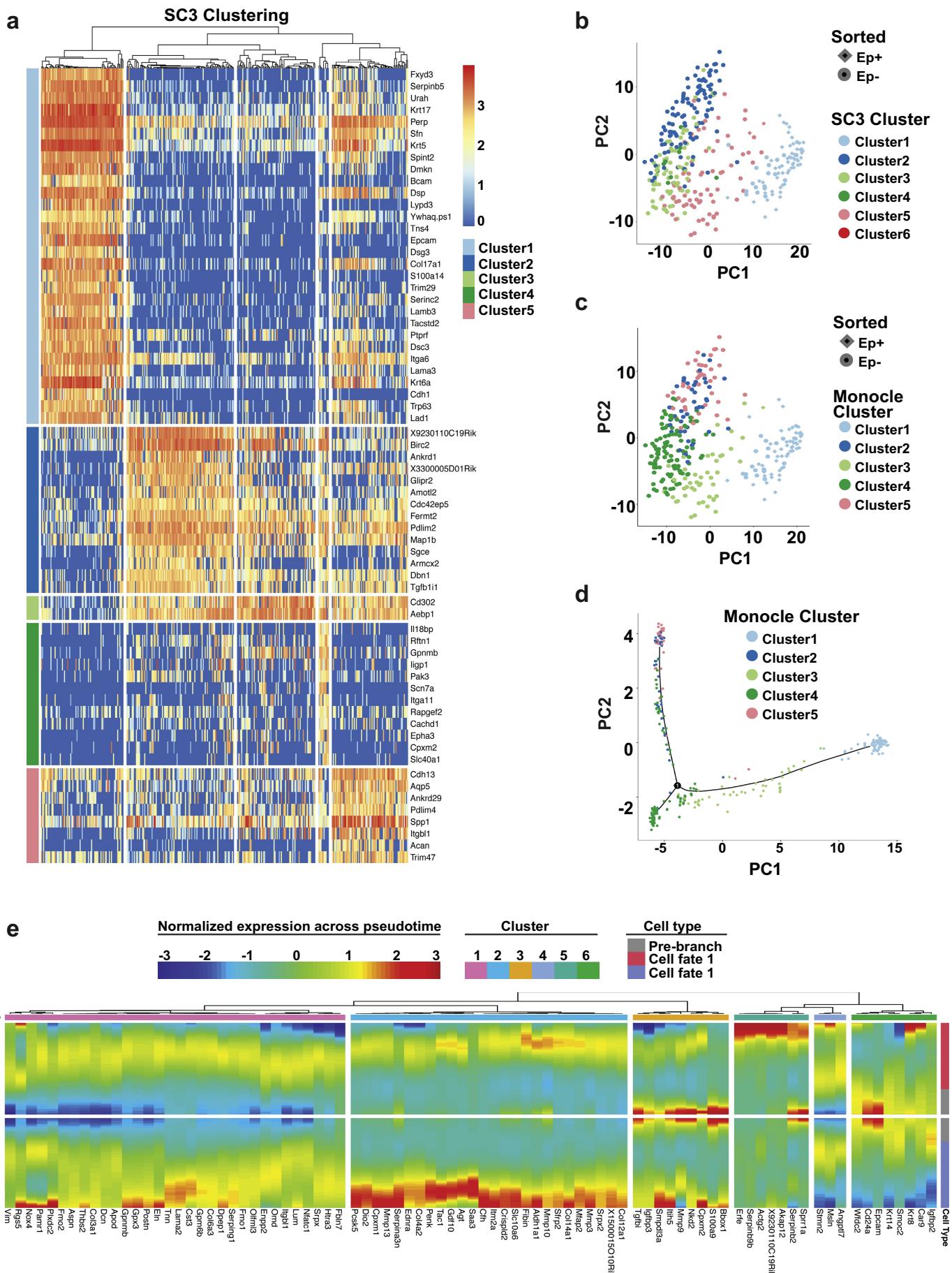
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Identification of EMT-related tumour heterogeneity by scRNA-seq. a, Heat map of normalized scRNA-seq expression of markers. Markers are obtained from the scRNA-seq data using SC3 with clustering parameter $k = 5$ and selecting genes with an area under the curve (AUC) > 0.8 and $P < 0.01$. Columns represent cells and rows represent genes. Colouring represents the normalized expression in the scRNA-seq data for each cell. **b**, Dimensionality reduction of scRNA-seq data using PCA colouring for clusters identified by SC3 clustering. Dots represent single cells; shape represents FACS sorting criteria. **c**, Dimensionality reduction of scRNA-seq data using PCA colouring for clusters identified by Monocle using only the top 500 genes with the highest covariability with CD51, CD61 and CD106 expression. Dots represent single cells; colour represents Monocle cluster; shape represents FACS sorting criteria. **d**, Pseudotime-ordering trajectory of scRNA-seq data using Monocle. The top 500 differentially expressed genes across clusters identified in a semi-supervised approach using Monocle were used to order cells in pseudotime using the DDRTree method. Dots represent single cells; colours represent assigned Monocle clusters. **e**, Heat map representing normalized scRNA-seq data after branched expression analysis modelling (BEAM) using Monocle for the top 100 differentially expressed genes across the branch-point. The pre-branch here is set to be the state corresponding to the Epcam$^+$ FACS-sorted cells. Cell fates 1 and 2 represent the two branches of the pseudotime trajectory corresponding to more hybrid and mesenchymal EMT phases, respectively. Heat map colours represent gene-wise normalized expression across pseudotime. Genes are hierarchically clustered according to their expression pattern. **a–e**, $n = 66$ Epcam$^+$ and $n = 277$ Epcam$^-$ cells from one tumour.
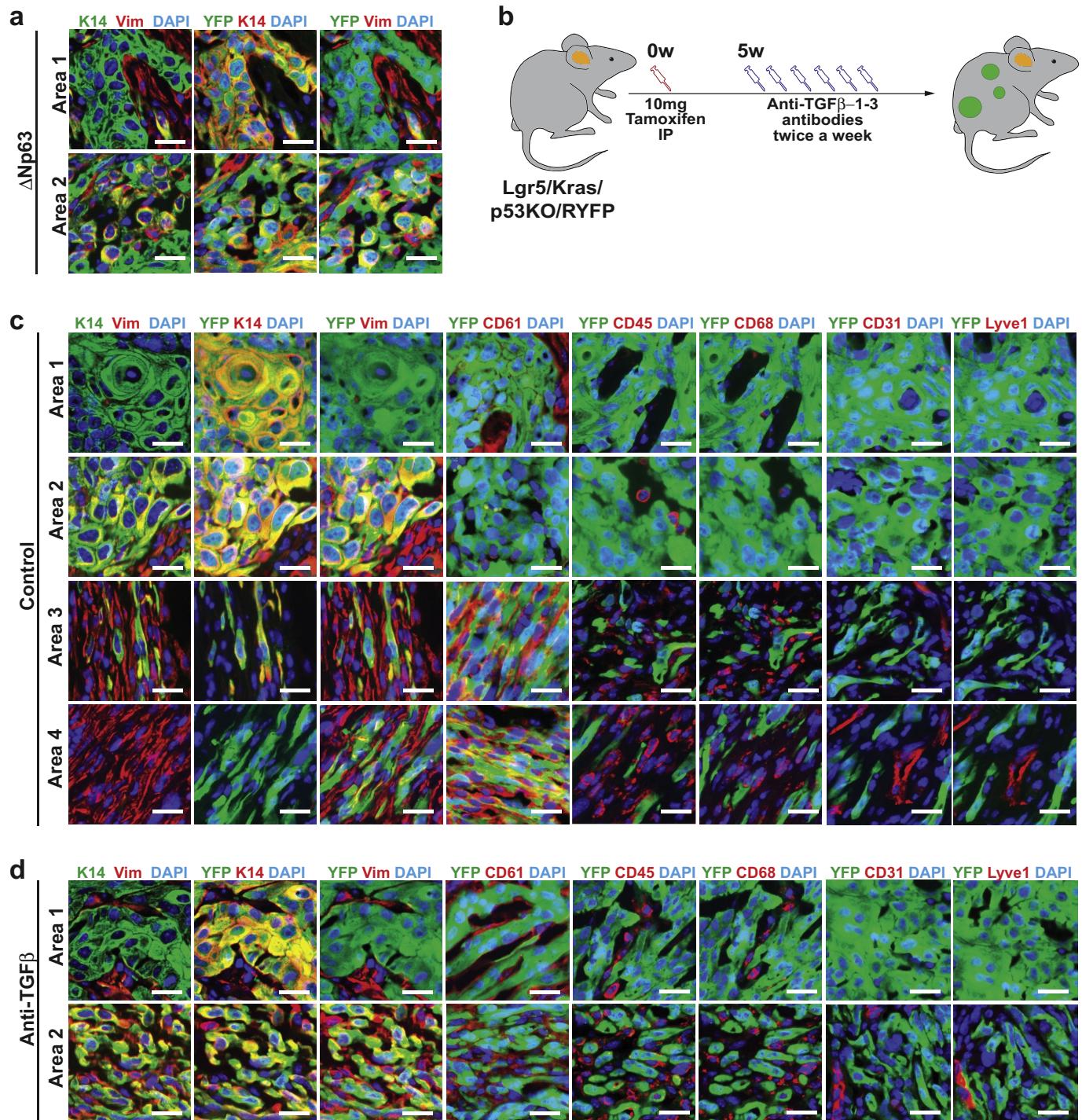
Extended Data Fig. 9 | Effect of constitutive overexpression of p63 and inhibition of TGF-β signalling in vivo on the different EMT subpopulations. a, Immunostaining of K14 and vimentin, YFP and K14 or YFP and vimentin in hair follicle-derived skin SCCs with constitutive overexpression of p63, showing that EMT is blocked at an early stage and only area 1 epithelial TCs (K14$^+$vimentin$^-$) and area 2 early hybrid TCs (K14$^+$vimentin$^+$) can be detected. Scale bars, 50 μm, $n = 5$ biological replicates. **b**, Experimental strategy to inhibit TGF-β signalling during hair follicle-derived SCC tumorigenesis. **c**, Immunostaining of K14 and vimentin or YFP K14, vimentin, CD61, CD45, CD68, CD31 or Lyve1 in area 1 (K14$^+$vimentin$^-$), area 2 (K14$^+$vimentin$^+$), area 3 (K14$^+$vimentin$^+$CD61$^+$) and area 4 (K14$^-$vimentin$^+$CD61$^+$) in mice treated with control antibodies. **d**, Immunostaining showing the expression of K14 and vimentin, or YFP and K14, vimentin, CD61, CD45, CD68, CD31 or Lyve1 in area 1 and area 2 of mice treated with anti-TGFβ antibodies, showing that inhibition of TGF-β or Smad2 signalling blocks EMT progression at the early stages of EMT and no areas corresponding to late EMT states (area 3 and 4) could be detected. **c, d**, Scale bars, 50 μm, $n = 3$ biological replicates.
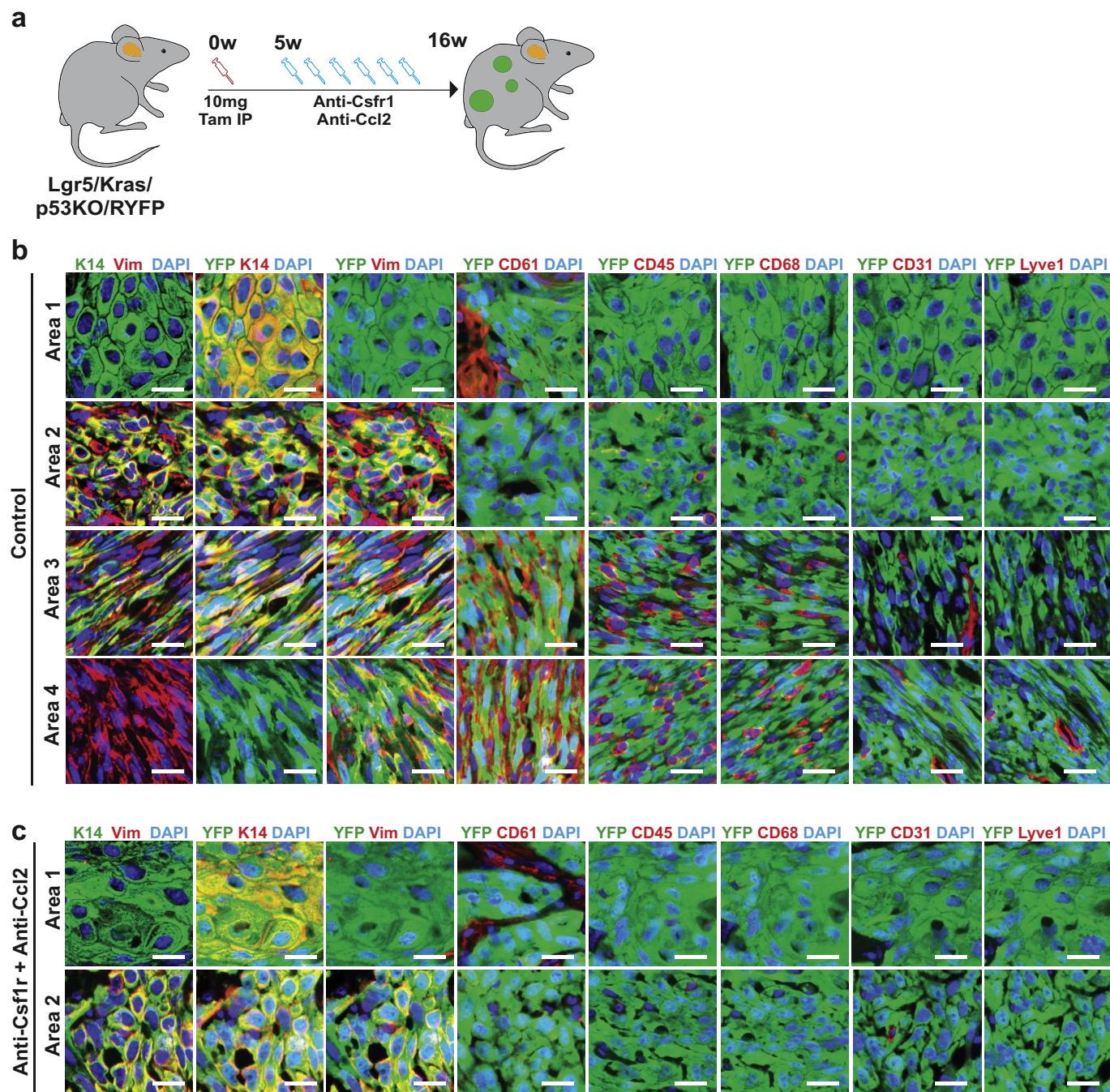
Extended Data Fig. 10 | Macrophage depletion inhibits EMT progression. a, Experimental strategy to deplete macrophages by administrating anti-Csfr1 and anti-Ccl2 antibodies. **b**, Immunostaining of K14 and vimentin or YFP and K14, vimentin, YFP, CD45, CD68, CD31 or Lyve1 in area 1 (K14$^+$vimentin$^-$), area 2 (K14$^+$vimentin$^+$), area 3 (K14$^+$vimentin$^+$CD61$^+$) and area 4 (K14$^-$vimentin$^+$CD61$^+$) in mice treated with control antibodies. Scale bars, 50 μm. **c**, Immunostaining of K14 and vimentin or YFP and K14, vimentin, CD61, CD45, CD68, CD31 or Lyve1 in area 1 and area 2 of mice treated with anti-Csf1r and anti-Ccl2 antibodies, showing that EMT is blocked at its early stages and no areas corresponding to late EMT stages (areas 3 and 4) could be detected. Scale bars, 50 μm, $n = 3$ biological replicates.

LETTER

A massive core for a cluster of galaxies at a redshift of 4.3

T. B. Miller[1,2]*, S. C. Chapman[1,3,4], M. Aravena[5], M. L. N. Ashby[6], C. C. Hayward[6,7], J. D. Vieira[8], A. Weiß[9], A. Babul[10], M. Béthermin[11], C. M. Bradford[12,13], M. Brodwin[14], J. E. Carlstrom[15,16,17,18], Chian-Chou Chen[19], D. J. M. Cunningham[1,20], C. De Breuck[19], A. H. Gonzalez[21], T. R. Greve[22], J. Harnett[23], Y. Hezaveh[24], K. Lacaille[1,25], K. C. Litke[26], J. Ma[21], M. Malkan[27], D. P. Marrone[26], W. Morningstar[24], E. J. Murphy[28], D. Narayanan[21], E. Pass[1,29], R. Perry[1], K. A. Phadke[8], D. Rennehan[10], K. M. Rotermund[1], J. Simpson[30,31], J. S. Spilker[26], J. Sreevani[8], A. A. Stark[6], M. L. Strandet[9,32] & A. L. Strom[33]

Massive galaxy clusters have been found that date to times as early as three billion years after the Big Bang, containing stars that formed at even earlier epochs[1-3]. The high-redshift progenitors of these galaxy clusters—termed 'protoclusters'—can be identified in cosmological simulations that have the highest overdensities (greater-than-average densities) of dark matter[4-6]. Protoclusters are expected to contain extremely massive galaxies that can be observed as luminous starbursts[7]. However, recent detections of possible protoclusters hosting such starbursts[8-11] do not support the kind of rapid cluster-core formation expected from simulations[12]: the structures observed contain only a handful of starbursting galaxies spread throughout a broad region, with poor evidence for eventual collapse into a protocluster. Here we report observations of carbon monoxide and ionized carbon emission from the source SPT2349-56. We find that this source consists of at least 14 gas-rich galaxies, all lying at redshifts of 4.31. We demonstrate that each of these galaxies is forming stars between 50 and 1,000 times more quickly than our own Milky Way, and that all are located within a projected region that is only around 130 kiloparsecs in diameter. This galaxy surface density is more than ten times the average blank-field value (integrated over all redshifts), and more than 1,000 times the average field volume density. The velocity dispersion (approximately 410 kilometres per second) of these galaxies and the enormous gas and star-formation densities suggest that this system represents the core of a cluster of galaxies that was already at an advanced stage of formation when the Universe was only 1.4 billion years old. A comparison with other known protoclusters at high redshifts shows that SPT2349-56 could be building one of the most massive structures in the Universe today.

In a multiband survey over 2,500 deg^2 of sky, the South Pole Telescope (SPT) discovered a population of rare ($n \approx 0.04\,deg^{-2}$), extremely bright (with observed flux densities (S) at 1.4 mm of more than 20 mJy), millimetre-selected sources[13,14]. Subsequent Atacama Large Millimeter/submillimeter Array (ALMA) imaging at wavelengths of 870 μm showed that more than 90% of these SPT-selected sources are single high-redshift submillimetre galaxies (SMGs)[15] with intrinsic flux densities of $S_{870\mu m} = 5$–10 mJy, gravitationally lensed by factors of 5–20 (ref. [16]), and with a median redshift of $z \approx 4$ (ref. [17]). However,

about 10% of these sources show no evidence of lensing and may instead be intrinsically very luminous galaxies or even groups of many rapidly star-forming galaxies. We made observations at 870 μm using a low-resolution bolometer camera on the APEX telescope (LABOCA), revealing the brightest such source in the SPT 2,500 deg^2 survey— SPT2349-56 ($S_{1.4mm} = 23.3$ mJy)—to consist of two elongated sources with a combined flux density of $S_{870\mu m} \approx 110$ mJy (Fig. 1), with the brighter, southern source comprising about 77 mJy of this flux density. A redshift survey by ALMA[17] further resolved SPT2349-56 into a pair of bright 3-mm sources associated with the southern LABOCA source, with both members of the pair lying at $z = 4.3$.

To better understand the nature of this structure, we undertook deep ALMA spectral imaging of the brighter southern peak of the extended LABOCA source. We used a 358-GHz map containing the redshifted [C II]$_{1900.5GHz}$ line to search for line-emitting galaxies. We carried out a blind spectral-line survey (see Methods) of the data cube, revealing 14 line emitters at redshifts of about 4.31 with high significance (with a signal-to-noise ratio of more than 7). We detected 12 of these emitters individually in the 1.1-mm continuum map at a significance of more than 5σ, with 1.1-mm flux densities ranging from 0.2 mJy to 5 mJy (Fig. 1). The remaining two line emitters (M and N) are both detected at lower significance in the 1.1-mm continuum map, but have robust counterparts detected by the Infrared Array Camera (IRAC) on board the Spitzer Space Telescope (Extended Data Table 1 and Extended Data Fig. 1). Eight of these sources are also detected (at significances of more than 5σ) in the carbon monoxide (4–3) line. The ALMA spectra are shown in Fig. 1.

Measurements of both the continuum and the spectral lines of the 14 galaxies allow us to estimate their star-formation rates (SFRs) and gas masses (Table 1 and Extended Data Table 1). The physical properties of these sources indicate that this protocluster already harbours massive galaxies that are rapidly forming stars from an abundant gas supply. The two brightest sources, A and B, have SFRs in excess of 1,000 Solar masses per year ($M_\odot\,yr^{-1}$) within their resolved, roughly 3-kpc radii. The total SFR of the 14 sources is $6,000 \pm 600\,M_\odot\,yr^{-1}$. Multicolour imaging with the Spectral and Photometric Imaging Receiver (SPIRE) instrument on board the Herschel Space Observatory (at wavelengths of 250 μm, 350 μm and 500 μm), in addition to the 870-μm LABOCA

Fig. 1 | The SPT2349-56 field and spectra of the constituent galaxies.
a, LABOCA 870-μm contours of SPT2349-56, overlaid on the IRAC 3.6-μm image; the 26″ beam at 870 μm is shown in white. Contours represent signal-to-noise rations of 3,7 and 9, moving inwards. The small red circles show the locations of the 14 protocluster sources.
b, ALMA band-7 imaging (276 GHz, 1.1 mm), showing the 14 confirmed protocluster sources, labelled A to N. Black and blue contours denote 75% (outer contour) and 90% (inner contour) of the peak flux for each source, based on the CO(4–3) and [C II] lines, respectively. The dashed black line shows where the primary beam is at 50% of its maximum. The filled blue ellipse shows the 0.4″ naturally weighted synthesized beam. **c**, CO(4–3) spectra (black lines) and [C II] spectra (yellow bars) for all 14 sources, centred at the biweight cluster redshift $z = 4.304$. The [C II] spectra are scaled down in flux by a factor of ten, for clarity of presentation. The red arrows show the velocity offsets determined by fitting a Gaussian profile to the CO(4–3) spectra for all sources except for G, H, K, L, M and N, for which we used [C II] (because these sources are not detected in the CO(4–3) spectra).

map, suggest that the northern LABOCA structure also lies at $z = 4.3$ (see Methods). The sources detected in the ALMA 870-μm imaging therefore comprise just 50% of the total flux density of the southern LABOCA source, and 36% of the total LABOCA flux density, suggesting that the roughly 500-kpc extent of the protocluster contains a total star-formation rate of $16,500\,M_\odot\,\mathrm{yr}^{-1}$. Modelling the spectral energy distribution on the basis of this combined submillimetre photometry yields an infrared luminosity (at wavelengths from 8 μm to 1,100 μm) of $(8.0 \pm 1.0) \times 10^{13}$ times the Solar luminosity ($L_\odot$). The gas masses of the 14 protocluster galaxies—estimated from the CO(4–3) line, or from the [C II] line if undetected in CO(4–3) (see Methods)—range from

$1 \times 10^{10}M_\odot$ to $1 \times 10^{11}M_\odot$, with a total gas mass of roughly 6×10^{11} ($X_{CO}/0.8)M_\odot$ (where X_{CO} is the conversion factor from CO(1–0) luminosity to total gas mass). A follow-up survey of colder molecular gas in the CO(2–1) line with the Australia Telescope Compact Array (ATCA, a radio telescope) detects the bulk of this large gas repository, especially in the central region near sources B, C and G, and confirms that the assumed line-intensity ratio, CO(4–3) to CO(1–0), used in the Methods when calculating the total gas mass, is consistent with the average measurements from ATCA.

The detected ALMA sources also enable an initial estimate of the mass of the protocluster. We determine the mean redshift using the biweight estimator[18] to be $\langle z \rangle_{bi} = 4.3040^{+0.0020}_{-0.0019}$. The velocity dispersion of the galaxy distribution is $\sigma_{bi} = 408^{+82}_{-56}$ km s^{-1} according to the biweight method[18], which is the standard approach for galaxy samples of this size. Other common methods (gapper[18] and Gaussian fit) agree to within 3% and provide similar errors. Under the assumption that SPT2349-56 is approximately virialized, the mass-dispersion relation for galaxy clusters[19] indicates a dynamical mass of $M_{dyn} = (1.16 \pm 0.70) \times 10^{13}M_\odot$, which is an upper limit if the system has not yet virialized. Given the possible selection effect of requiring a bright source (with $S_{1.4mm}$ values of more than 15 mJy) within the 1′ SPT beam for detection, we also further consider the possibility that our structure may represent an end-on filament being projected into a compact but unbound configuration, rather than a single gravitationally bound halo. Our analysis in the Methods suggests that this is not as likely as a relatively bound system in a massive halo, given the velocity dispersion measured as a function of position, and other supporting arguments. However, we cannot rule this possibility out completely, and further analysis and observations of the larger angular scale of the structure will be required to more fully understand the nature of this system.

Table 1 | Derived physical properties of SPT2349-56 protocluster members

Source	ΔV [km s^{-1}][†]	SFR [$M_\odot$ yr^{-1}]	M_{gas} [$10^{10}M_\odot$]
A	-90 ± 35	$1,170 \pm 390$	12.0 ± 2.1
B	-124 ± 31	$1,227 \pm 409$	11.2 ± 2.0
C	603 ± 12	907 ± 302	6.7 ± 1.2
D	-33 ± 40	530 ± 182	8.4 ± 1.5
E	84 ± 21	497 ± 179	4.8 ± 0.9
F	395 ± 82	505 ± 169	3.4 ± 0.7
G	308 ± 42	409 ± 137	2.9 ± 1.3‡
H	-719 ± 28	310 ± 105	4.4 ± 2.0‡
I	310 ± 78	268 ± 91	2.2 ± 0.5
J	-481 ± 35	243 ± 85	2.2 ± 0.5
K	631 ± 12	208 ± 71	3.1 ± 1.4‡
L	-379 ± 18	122 ± 43	3.3 ± 1.5‡
M	34 ± 21	75 ± 34	1.2 ± 0.6†
N	90 ± 25	64 ± 29	1.0 ± 0.5†

†Velocity offsets were measured relative to the mean redshift, $z = 4.304$.
‡The [C II] line was used to derive M_{gas} in these cases, as CO(4–3) was not detected.
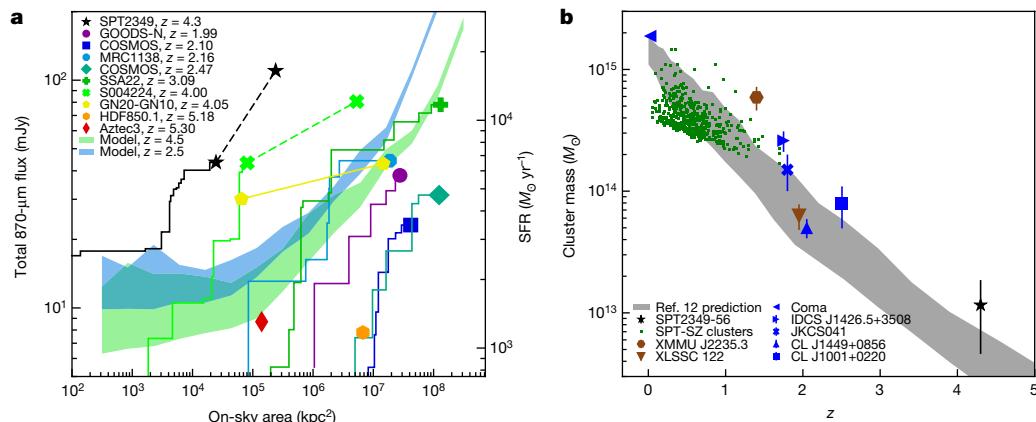
Fig. 2 | Comparison of SPT2349-56 to other cluster and protocluster systems. a, The cumulative 870-μm flux density is plotted against on-sky area for SPT2349-56 and for other SMG-rich overdensities at high redshifts (see Methods). The solid black line shows the ALMA-identified sources in SPT2349-56, while the dashed line includes the wider-field LABOCA-detected structure. The blue and green shaded regions denote the maximum flux density versus area curves obtained in 100 Monte Carlo realizations of a theoretical model for submillimetre-luminous protoclusters at $z = 4.5$ and $z = 2.5$, based on an N-body simulation (see Methods). Most of the literature SMG overdensities are consistent with the model expectations, whereas SPT2349-56 lies vastly above the region spanned by the model. A recently discovered $z = 4$ protocluster[23], SMM J004224 (light green cross), is quite a unique system, but more

than ten times less dense than—and probably only about 50% the total luminosity of—SPT2349-56. **b**, Cluster mass versus redshift is shown for SPT2349-56 and for other massive galaxy clusters (from the literature) that have detected intracluster media and well defined masses. The colour scheme highlights the different methods used for selecting massive clusters (brown, X-ray; blue, optical; green, Sunyaev–Zel'dovich effect). Error bars represent 1σ standard deviation. We also show the mean protocluster most-massive-progenitor mass versus redshift relation predicted by N-body simulations[12]. The location of SPT2349-56 in this plane suggests a very massive descendant (with a halo mass of more than $10^{15} M_\odot$ at $z = 0$), although we caution that the complex growth histories of dark-matter halos make it difficult to reliably predict the $z = 0$ halo mass from the halo mass at a given epoch[20].

If the total halo mass represented by these 14 SMGs is indeed roughly $10^{13} M_\odot$, then the protocluster could conceivably be the progenitor of a galaxy cluster that is larger than $10^{15} M_\odot$, comparable to the Coma cluster at $z = 0$ (Fig. 2), as we deduce from simulations that track the hierarchical halo growth of an approximately $10^{15} M_\odot$ galaxy cluster from early epochs[12]. The location of SPT2349-56 in this particular plane suggests a very massive descendant, but we caution that N-body simulations indicate that it is difficult to reliably predict $z = 0$ halo masses from the halo mass at a given epoch, owing to the large halo-to-halo variations in dark-matter halo-growth histories[20].

To study the relative overdensity and concentration of SPT2349-56, it is desirable to compare it with other active protoclusters at high redshifts. SPT2349-56 is highly overdense, as it harbours ten SMGs with $S_{1.1mm}$ values of more than about 0.5 (a level at which we are 'complete', meaning that we are not missing any sources and that there is uniform sensitivity across our search area) located within a circle of diameter 19″ (130 kpc), corresponding to a number density of $N(S_{1.1mm} > 0.5\,\mathrm{mJy}) \approx 2 \times 10^4\,\mathrm{deg}^{-2}$. By comparison, the average number of field sources with $S_{1.1mm}$ values of more than 0.5 mJy within this area across all redshifts is less than one[21]; thus, the SPT2349-56 field is overdense by more than a factor of ten. When we account for the fact that all sources in the SPT2349-56 field are at the same redshift, the volume density is more than 1,000 times the field density, assuming a redshift binning of $\Delta z = 0.1$ and the redshift distribution for SMGs[22]. In Fig. 2, we plot 'curves of growth' of the total 870-μm flux density against on-sky area for SPT2349-56 and other SMG-rich protoclusters (see Methods for details of the comparison sample). For SPT2349-56, we plot both the total flux density of the 14 confirmed protocluster members detected with ALMA, and the total flux density of the extended LABOCA structure. The curve of growth for SPT2349-56 rises much more steeply than those of the other high-redshift protoclusters, demonstrating its extreme density. For SPT2349-56, the on-sky area encompassing the accumulated 870-μm flux density (and thus approximately the total SFR) is as much as three orders of magnitude lower than for other protoclusters at $z > 2$. Thus SPT2349-56 clearly stands out as the densest collection of SMGs: although some other protoclusters contain as many SMGs, they extend over much larger areas on the sky, with separations often exceeding 10 arcmin (22 co-moving megaparsecs (cMpc) to

15 cMpc at redshifts of 4.3 to 2). This comparison shows that SPT2349-56 has probably been observed during a much more advanced stage of cluster formation than other high-redshift protoclusters—as a cluster core that is in the process of assembly, rather than an extended structure that may not even collapse to form a cluster by the present day[12].

Also shown in Fig. 2 is the maximal curve of growth predicted by a theoretical model for submillimetre-luminous protocluster regions at redshifts of about 4.5 (see Methods). Except for SPT2349-56 and the recent Herschel discovery SMM J004224 (ref. [23]), the comparison high-redshift protoclusters exhibit $S_{870\mu m}$ curves of growth that are fairly consistent with the model expectations. The model prediction for the region spanned by SPT2349-56 is roughly 10% of the observed total flux density of the 14 ALMA sources. The underprediction is more severe if we consider the extended LABOCA source: only about 5% of the observed flux density is recovered. This discrepancy may suggest that environmental effects (such as enhanced galaxy interactions or gas accretion in high-density environments) that are not included in this theoretical model are responsible for the extremely high SFR density exhibited by SPT2349-56. An alternative theoretical approach—'zoom' hydrodynamical simulations of protoclusters[24]—can potentially capture such environmental effects. But so far such simulations have been unable to reproduce the extremely high SFR inferred for SPT2349-56: of the 24 protocluster simulations presented by these authors[24], the maximum total SFR attained was about 1,700 $M_\odot$ yr^{-1}, an order of magnitude less than that of SPT2349-56. However, the volume of the N-body simulation from which the 24 halos were selected was $1\,h^{-3}\,\mathrm{cGpc}^3$, which may be too small to contain an object as rare as SPT2349-56. Nevertheless, the existence of SPT2349-56, which contains an unprecedented concentration of rapidly star-forming SMGs from a time when the Universe was only 1.4 billion years old, poses a formidable challenge to theoretical models seeking to explain the origin and evolution of galaxy (proto)clusters.

As outlined above, SPT2349-56 may represent a much more advanced stage of cluster formation than the typical $z > 4$ protoclusters identified to date. Given that the cores of present-day galaxy clusters are characterized by massive elliptical galaxies with old-to-intermediate-age stellar populations[25], and that SMGs are thought to be the high-redshift progenitors of present-day ellipticals[22], it is likely that at least

some of the 14 SMGs located at the same redshift within a region of less than 130 kpc in diameter will soon merge to form a massive elliptical galaxy at the core of a lower-redshift galaxy cluster.

Theoretical studies have shown that at redshifts of more than 4, the progenitors of galaxy clusters should span more than 5 cMpc (refs. [12,26]), corresponding to an angular scale of as much as a degree; we are thus possibly observing only a small part of a much larger structure. For SPT2349-56, it is unknown whether the overdensity extends over such a large scale, as more detailed observations are required to characterize the field surrounding SPT2349-56. We have shown that the extended LABOCA-detected complex has submillimetre colours similar to the core region identified by our ALMA observations, and thus it is likely that the entire complex is found at redshifts of about 4.3. We have also identified five additional bright SPIRE sources in the surrounding roughly 18×18 cMpc field, with similar red colours lying several arcmin from the core structure (see Methods). These are candidates for further protocluster members located in an extended, collapsing structure, similar to the comparison SMG overdensities shown in Fig. 2. If all of these sources are confirmed to lie at $z = 4.31$, this would approximately double the far-infrared luminosity of the cluster, making it by far the most active system known in the Universe. Given that SPT2349-56 was selected from a blind millimetre survey of 2,500 deg^2 (approximately one-sixteenth of the sky), it is unlikely there are more than about 16 such structures across the entire sky. A full analysis of other unlensed sources from the SPT survey, to identify possible systems similar to SPT2349-56, will place stronger constraints on early structure formation in the Universe.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0025-2.

1. Wang, T. et al. Discovery of a galaxy cluster with a violently starbursting core at $z = 2.506$. *Astrophys. J.* **828**, 56 (2016).
2. Mantz, A. B. et al. The XXL Survey: XVII. X-ray and Sunyaev-Zel'dovich properties of the redshift 2.0 galaxy cluster XLSSC 122. *Astron. Astrophys.* https://doi.org/10.1051/0004-6361/201630096 (2017).
3. Stanford, S. A. et al. IDCS J1426.5+3508: discovery of a massive, infrared-selected galaxy cluster at $z = 1.75$. *Astrophys. J.* **753**, 164 (2012).
4. Springel, V. et al. Simulating the joint evolution of quasars, galaxies and their large-scale distribution. *Nature* **435**, 629–636 (2005).
5. Overzier, R. A. et al. ΛCDM predictions for galaxy protoclusters—I. The relation between galaxies, protoclusters and quasars at $z \sim 6$. *Mon. Not. R. Astron. Soc.* **394**, 577–594 (2009).
6. Chiang, Y.-K., Overzier, R. A., Gebhardt, K. & Henriques, B. Galaxy protoclusters as drivers of cosmic star formation history in the first 2 Gyr. *Astrophys. J. Lett.* **844**, L23 (2017).
7. Miley, G. & De Breuck, C. Distant radio galaxies and their environments. *Astron. Astrophys. Rev.* **15**, 67–144 (2008).
8. Casey, C. M. et al. A massive, distant proto-cluster at $z = 2.47$ caught in a phase of rapid formation? *Astrophys. J. Lett.* **808**, L33 (2015).
9. Chapman, S. C. et al. Do submillimeter galaxies really trace the most massive dark-matter halos? Discovery of a high-z cluster in a highly active phase of evolution. *Astrophys. J.* **691**, 560–568 (2009).
10. Tamura, Y. et al. Spatial correlation between submillimetre and Lyman-α galaxies in the SSA22 protocluster. *Nature* **459**, 61–63 (2009).
11. Ma, J. et al. Stellar masses and star formation rates of lensed, dusty, star-forming galaxies from the SPT survey. *Astrophys. J.* **812**, 88 (2015).
12. Chiang, Y.-K., Overzier, R. & Gebhardt, K. Ancient light from young cosmic cities: physical and observational signatures of galaxy proto-clusters. *Astrophys. J.* **779**, 127 (2013).
13. Vieira, J. D. et al. Extragalactic millimeter-wave sources in the South Pole Telescope survey data: source counts, catalog, and statistics for an 87 square-degree field. *Astrophys. J.* **719**, 763–783 (2010).
14. Mocanu, L. M. et al. Extragalactic millimeter-wave point source catalog, number counts and statistics from 771 square degrees of the SPT-SZ survey. *Astrophys. J.* **779**, 61 (2013).
15. Vieira, J. D. et al. Dusty starburst galaxies in the early Universe as revealed by gravitational lensing. *Nature* **495**, 344–347 (2013).
16. Spilker, J. et al. ALMA imaging and gravitational lens models of South Pole Telescope-selected dusty, star-forming galaxies at high redshifts. *Astrophys. J.* **826**, 112 (2016).
17. Strandet, M. L. et al. The redshift distribution of dusty star forming galaxies from the SPT survey. *Astrophys. J.* **822**, 80 (2016).
18. Beers, T. C. et al. Measures of location and scale for velocities in clusters of galaxies—a robust approach. *Astron. J.* **100**, 32 (1990).
19. Evrard, A. E. et al. Virial scaling of massive dark matter halos: why clusters prefer a high normalization cosmology. *Astrophys. J.* **672**, 122–137 (2008).
20. Cole, S., Helly, J., Frenk, C. S. & Parkinson, H. The statistical properties of Lambda cold dark matter halo formation. *Mon. Not. R. Astron. Soc.* **383**, 546–556 (2008).
21. Ono, Y., Ouchi, M., Kurono, Y. & Momose, R. Faint submillimeter galaxies revealed by multifield deep ALMA observation: number counts, spatial clustering, and a dark submillimeter line emitter. *Astrophys. J.* **795**, 5 (2014).
22. Simpson, J. M. et al. An ALMA survey of submillimeter galaxies in the extended Chandra deep field south: the redshift distribution and evolution of submillimeter galaxies. *Astrophys. J.* **788**, 125 (2014).
23. Oteo, I. et al. An extreme proto-cluster of luminous dusty starbursts in the early Universe. Preprint at https://export.arxiv.org/pdf/1709.02809 (2017).
24. Granato, G. L. et al. The early phases of galaxy clusters formation in IR: coupling hydrodynamical simulations with GRASIL-3D. *Mon. Not. R. Astron. Soc.* **450**, 1320–1332 (2015).
25. Renzini, A. Stellar population diagnostics of elliptical galaxy formation. *Annu. Rev. Astron. Astrophys.* **44**, 141–192 (2006).
26. Oñorbe, J. et al. How to zoom: bias, contamination and Lagrange volumes in multimass cosmological simulations. *Mon. Not. R. Astron. Soc.* **437**, 1894–1908 (2013).

METHODS

Observations. *SPT and SPIRE discovery*. The SPT[27] possesses a unique combination of sensitivity, selection wavelengths (3 mm, 2 mm and 1.4 mm) and beam size that potentially make it ideal for finding the active core regions of galaxy clusters forming at the earliest epochs. Finding very distant ($z > 4$), gravitationally lensed millimetre sources in the SPT survey is quite straightforward, as the contrast to such distant bright sources is high relative to the weak (generally undetected) galactic foregrounds (Extended Data Fig. 2). However, searching for the rare SMGs in the SPT 2,500 deg^2 survey that are unlensed—and therefore are candidates for active groups and protoclusters like SPT2349-56—involves sifting through the many gravitationally lensed sources, and typically involves multistage follow-up efforts using various facilities: a single-dish mapping instrument such as APEX-LABOCA to better localize the emission within the roughly 1′ SPT beam; deep optical imaging to search for bright lensing galaxies; and high-resolution ALMA mapping. The spatially extended sources in SPT2349-56 found with LABOCA span more than an arcmin. With the upcoming deep surveys using the next-generation SPT-3G receiver, this 'extended-beam' thermal-source structure may present a unique signature of many early forming protoclusters, affording the first complete census of the early epochs of structure formation.

A shallow, wide-field image taken by SPIRE (on board the Herschel observatory) over a 100 deg^2 subregion of the SPT-SZ survey[28] reveals the red colours of SPT2349-56, and also that SPT2349-56 appears to reside in something of a void in the $z \approx 1$ foreground that dominates the SPIRE galaxy population. However, the high redshift of SPT2349-56 means that it is not much brighter than many other SPIRE sources in this field, and, aside from its colours, SPT2349-56 does not stand out substantially from the field despite its extreme properties. SPT2349-56 is not detected in the all-sky Planck survey[29], the lower sensitivity of Planck compared with SPT being exacerbated by beam dilution in the 3′ Planck beam.

Obtaining the redshift for SPT2349-56 was beyond the scope of the original SPT–SubMillimeter Galaxy (SMG) redshift survey, owing to the faintness of the unlensed components relative to the typical bright, gravitationally lensed SMGs found in the bulk of the SPT-SMG sample. In ALMA channels Cy 0 and 1, SPT2349-56 was included in the 3-mm spectral-scan redshift survey[15,30], but no lines were detected in the short, roughly 1-min integrations with about 16 ALMA antennae. In Cy 3, a deeper follow-up 3-mm spectral scan was able to tentatively identify two CO lines and a double-source structure with a likely redshift of $z = 4.30$, confirmed by APEX/FLASH C+ detection[17].

APEX–LABOCA discovery. We obtained 870-μm imaging of SPT2349-56 using LABOCA on the APEX telescope. A shallow image with a 1.6-hour integration time was observed on 27 Sep 2010, reaching a noise level of about 5 mJy per beam (root mean square, r.m.s.). In August 2017 we obtained a deeper image (18.8-hour integration time; project ID E-299.A-5045A-2017; principal investigator S.C.C.), reaching a minimum noise level of 1.3 mJy per beam, and less than 2.0 mJy per beam r.m.s. for 75.3 arcmin^2 and less than 1.5 mJy per beam r.m.s. for 32.4 arcmin^2 (shown in Fig. 1 and Extended Data Fig. 3). All observations were carried out using standard raster-spiral observations[31] under good weather conditions (precipitable water vapour was 0.6 mm and 0.8 mm for the 2010 and 2017 observing campaigns, respectively). Calibration was achieved through observations of Uranus, Neptune and secondary calibrators, and was found to be accurate within 8.5% r.m.s. The atmospheric attenuation was determined via skydips every 2 h as well as from independent data from the APEX radiometer, which measures the line-of-sight water vapour column every minute. The data were reduced and imaged using the BoA reduction package[32]. LABOCA's central frequency and beam size are 345 GHz and 19.2″, resolving the SPT 1.4-mm elongated source into two bright LABOCA sources.

Both LABOCA observations yield consistent calibration results, with a peak intensity at 21″ resolution of 50 mJy per beam for the brighter, southern component (right ascension (RA) 23 h 49 min 42.70 s; declination (dec.) −56° 38′ 23.4″). In addition the LABOCA map reveals a second source to the north at RA 23 h 49 min 42.86 s, dec. −56° 37′ 31.02″, with a peak flux density at 21″ resolution of 17 mJy per beam. Both sources are clearly extended even at LABOCA's relatively coarse spatial resolution, with an observed source size of 28″ × 25″ and 31″ × 24″ for the southern and northern sources, respectively. These components are connected by a faint bridge emission. The total 870-μm flux density of the SPT2349-56 system is 110.0 ± 9.5 mJy, of which about 77 mJy are associated with the southern component, 25 mJy with the northern component, and 7 mJy with the connection between the components (using the subregions shown in Extended Data Fig. 3). One additional submillimetre source is detected at >5σ in the LABOCA image to the east of the primary source, but has blue colours inconsistent with $z \approx 4$, and is not likely to be a member of the extended protocluster.

ALMA imaging and spectroscopy. Observations made using ALMA band 3 targeted the CO(4–3) line in SPT2349-56, centred in the lowest frequency of the spectral windows adopted (86–88 GHz), taken under a cycle 3 program (2015.1.01543.T; principal investigator K.L.). Data were taken on 24 June 2016 with a 47-min

integration time. The array used 36 antennas with baselines ranging from 15 m to 704 m, and provided a naturally weighted synthesized beam size of about 1″. The asteroid Pallas and source J2343-5626 were used to calibrate the flux and phase respectively. Data were processed using the standard ALMA pipeline with natural beam weighting.

ALMA band-7 imaging (276 GHz) was obtained under a cycle 4 program (2016.0.00236.T; principal investigator S.C.C.), targeting the peak of the brightest LABOCA source. Observations were obtained on 14 December 2016 in a 40–2 array configuration with baseline lengths of 15–459 m, giving a naturally weighted synthesized beam size of about 1″. There were 40 antennas available, with a total on-source integration time of 22 min. The asteroid Ceres and J2357-5311 were used as flux and phase calibrators respectively. The [C II] line ($\nu_{rest} = 1{,}900.5$ GHz) was observed as part of the same ALMA project on 23 March 2017, tuning in band 7 to the redshifted line at $\nu_{obs} = 358.3$ GHz in the upper sideband covering 356–360 GHz. These observations used the 40–2 array configuration with baselines of 16–459 m, giving a naturally weighted synthesized beam size of about 0.5″. An on-source integration time of 14 min was obtained, and J2357-5311 was used as both flux and phase calibrator. The data were reprocessed using the Common Astronomy Software Applications (CASA) and the standard ALMA-supplied calibration, using natural beam weighting to maximize sensitivity.

One-dimensional spectra were extracted from the centroid of the line emission for each source and binned into 75 km s^{-1} channels. Spectra are presented in Fig. 1, and are smoothed using a Gaussian filter with a full width at half-maximum (FWHM) of 100 km s^{-1} for presentation. A Gaussian line profile is fit using a least-squares method, providing errors to the velocity offsets from $z = 4.304$ in Table 1 and line widths in Extended Data Table 2. The continuum level is left as a free parameter in the fitting function, and is then subtracted to derive line fluxes and for presentation.

Blind search for [C II]. We performed a blind search for [C II] line emission in the ALMA band 7 data cube towards SPT2349-56. For this, we followed the procedure used to detect line emitters in the ASPECS survey[33]. We use a data cube channelized at 100 km s^{-1}, without primary beam correction and continuum subtraction. We used the Astronomical Image Processing System (AIPS) task SERCH. This task convolves the data cube along the frequency axis with a Gaussian kernel defined by different input linewidths, subtracts the surrounding continuum, and reports all channels and pixels that have a signal-to-noise ratio (SNR) over a specified limit. The SNR is defined as the maximum significance level achieved after convolving over the Gaussian kernels. We used a set of different Gaussian kernels, from 200 km s^{-1} to 600 km s^{-1}, and searched for all line peaks with SNRs of more than 4.0.

Once all peaks were identified, we used the IDL routine CLUMPFIND[34] to isolate individual candidates. A full list of 68 positive line peaks with SNRs of more than 4.0 was thus obtained. We quantified the reliability of our line search on the basis of the number of negative peaks in our ALMA cube, using the same line procedure. We found 43 negative peaks with SNRs of less than 5.8, and none at a higher SNR. This means that all positive line candidates with SNRs greater than 6.0 are probably real (100% purity). Out of the 14 [C II] line candidates detected, all have SNRs of more than 6.3, and 12 are associated with continuum detections in the ALMA data.

ATCA CO(2–1) line detection: observations. We used ATCA in its H168 hybrid array configuration to observe the CO(2–1) emission line ($\nu_{rest} = 230.5380$ GHz) of SPT2349-56 (with a primary beam size of 53″). The observations were performed as part of project C2818 during 2, 3 and 11 October 2016 under good weather conditions (atmospheric seeing values 90–400 m), and with five working antennae.

We used the ATCA 7-mm receivers, with the compact array broadband backend configured in the wide-bandwidth mode[35]. This leads to a total bandwidth of 2 GHz per correlator window and a spectral resolution of 1 MHz per channel (6.9 km s^{-1} per channel). The spectral windows were centred at observing frequencies of 43.5 GHz and 45.0 GHz, and aimed at observing the CO line and continuum emission, respectively.

Gain and pointing calibrations were performed every 10 min and 1 h, respectively. The bright sources 1921-293, 1934-638 and 2355-534 were used as bandpass, flux and gain calibrators, respectively. We expect the flux calibration to be accurate to within 15%, given the comparison of the Uranus and 1934-638 fluxes. The software package MIRIAD[36] and CASA[37] were used for editing, calibration and imaging.

The calibrated visibilities were inverted with the CASA task CLEAN using natural weighting. No cleaning was applied, given the relatively low significance of the CO-line detection in individual channels. The final data cube, averaged along the spectral axis, yields an r.m.s. of 0.23 mJy per beam for each 100 km s^{-1} channel, with a synthesized beam size of 5.6″ × 4.5″ (position angle 70.4°) at 43.5 GHz.

ATCA CO(2–1) line detection: results. One source was formally detected at the centre, corresponding to C II/continuum sources B + C + G. The central source (C) is unresolved at the resolution of the ATCA observations. The other two sources

are marginally detected to the west and north of the central source, coinciding with the location of C II/continuum sources D + E and A + K, respectively. We extracted spectra at these locations and obtained integrated line intensities by fitting Gaussian profiles to the identified line emission (Extended Data Fig. 4).

We computed CO luminosities using the integrated line intensities. We computed gas masses by assuming an ultraluminous infrared galaxy (ULRIG) X_{CO} factor of 0.8 ($M_{\odot}$ (K km s^{-1} pc^2)$^{-1}$), also assuming that the CO gas is in local thermodynamic equilibrium: thus $L'_{CO2-1} \approx L'_{CO1-0}$ (ref. [38]). The results of the CO-line observations are summarized in Extended Data Table 2 Collapsing the line-free spectral window along the spectral axis over the 2-GHz bandwidth leads to a non-detection of the continuum emission down to 80 μJy per beam (3σ standard deviation).

These results confirm the finding from CO(4–3) line that the main reservoir (72%) of molecular gas resides in the B + C + G system, with a smaller fraction hosted in the west and north locations.

Spitzer imaging. This field was twice observed at 3.6 μm and 4.5 μm with the IRAC[39] on board the Spitzer Space Telescope[40]. It was first observed in August 2009 as part of a large program to obtain follow-up imaging of a large sample of SPT-selected SMG sources (project 60194; principal investigator J.D.V.). The observing scheme for this project involved obtaining 36 dithered 100-s integrations at 3.6 μm and, separately, a much shallower 12 × 30-s integration at 4.5 μm. Later, in cycle 8, the field was covered serendipitously as part of the Spitzer–SPT deep-field survey (project 80032; principal investigator M.A. Stanford[40]). This project surveyed 92 deg^2 uniformly in both IRAC passbands, to a depth of 4 × 30 s. Using established techniques, we combined all exposures covering the SPT target from projects 60194 and 80032 at 3.6 μm and 4.6 μm, to obtain the best possible SNR in our final mosaics, which were pixellated to 0.5″. Nine of the 14 sources identified by ALMA are detected in the IRAC bands at greater than 3σ in at least one of the 3.6-μm or 4.5-μm channels (Extended Data Fig. 1).

Analysis of the surrounding field with SPIRE and LABOCA imaging. In Extended Data Fig. 5, our deep SPIRE RGB image is shown with LABOCA contours overlaid. A source sample was culled from the 250-μm-selected catalogue (there were 135 sources with SNR (250 μm) values of greater than 3 in an area of 52 arcmin^2), where the source peaks are best defined. To account for the large beam-size difference with SPIRE (ranging from 36″ at 500 μm to 18″ at 250 μm), we used a deblending code, with the 250-μm positions as spatial priors, which provides the standard parameters as well as the covariance matrices highlighting the degeneracies (almost none at 250 μm, but statistically significant at 500 μm). This code, FASTPHOT[41], takes into account these degeneracies to estimate the flux measurement errors.

Colour–colour and colour–flux diagrams are shown in Extended Data Fig. 5. The colour–colour diagram shows a 250-μm-selected sample with an SNR(250 μm) of more than 3, and is dominated by the $z \approx 1$ cosmic infrared background (blue and green colours) in the foreground of SPT2349-56. The colour–flux diagram shows an additional SNR(500 μm) of more than 3, cut to highlight just the well detected 500-μm source subsample. These diagrams highlight the extreme and red properties of SPT2349-56, but make clear that one of the three 250-μm peaks within the SPT2349-56 LABOCA structure is very likely to be a foreground galaxy (the green symbol highlighted in the figure shows very blue colours). Nevertheless, a full ALMA mapping of the structure is warranted given the uncertainties involved in the SPIRE deconvolution procedure.

Five red sources consistent with $z \approx 4$ ($S_{500μm} > S_{350μm} > S_{250μm}$; Extended Data Table 3) are found in the surrounding roughly 10′ × 10′ field, and are candidates for additional protocluster members in an extended, collapsing structure. If all these sources are bona fide $z = 4.3$ sources, this would substantially increase the total 870-μm flux density (and thus the far-infrared luminosity) of the cluster beyond the 110 mJy found in the central structure, making it by far the most active system known in the Universe (see Fig. 2).The deep LABOCA map marginally detects the closest of the five red SPIRE sources at about 3σ, consistent with expectations given the SPIRE flux densities. Full analysis of these surrounding SMGs will require follow-up efforts.

Properties, comparisons and simulations. *Derivation of physical properties.* We briefly describe here our procedures for calculating various physical quantities from the observables. To derive the star-formation rate, we measured the 870-μm flux density directly in the lower sideband (line-free bands) of our ALMA band-7 observations from cycle 4, finding consistent measurements with those in previous shallower observations[16]. We adopt an SFR-to-$S_{870μm}$ ratio of 150 ± 50 $M_{\odot}$ yr^{-1} mJy^{-1}, which is typical for SMGs[42]. The uncertainty in this ratio derives from variations in the dust temperature distribution among the SMG population, which are driven mainly by differences in the ratio of the luminosity absorbed by dust to the total dust mass[43]. This, combined with the measurement error, dictates the error on the SFR in Table 1.

We calculated the gas mass from the CO(4–3) line luminosity, which we converted to CO(1–0) luminosity by using a ratio between the brightness temperatures of these lines, $r_{4,1} = 0.41 \pm 0.07$, that was found from the average of a sample of

unlensed SMGs with multiple CO-line transitions[44]. We used a conservative conversion factor, $\alpha_{CO} = 0.8 \frac{M_{\odot}}{\text{K km s}^{-1} \text{pc}^2}$, and multiplied by 1.36 to account for the addition of helium. When CO(4–3) was not detected to a large extent, we used our [C II] line luminosity and the average CO(4–3)/[C II] ratio for our detected sample (denoted with double daggers in Table 1).

Spectral energy distribution of SPT2349-56. The SPT, LABOCA and SPIRE measurements resolve the SPT2349-56 structure to varying degrees, but none can isolate the core region resolved by our present ALMA observations with any confidence. We thus assembled a photometric catalogue of the total SPT2349-56 flux density from 250 μm to 850 μm, and modelled the resulting total spectral energy distribution (SED) to estimate some global properties of the system. Although the SPT does measure the flux at 1.4 mm, 2.0 mm and 3.0 mm (being 27.9 ± 4.6 mJy, 5.2 ± 1.1 mJy and 0.5 ± 0.1 mJy, respectively), we did not include these points in the SED fit, because the measurements are uncertain owing to the elongated structure of SPT2349-56 and difficulties with the filtering used to make the map. At IRAC wavelengths in the mid-infrared, we identified nine SMGs detected at more than 3σ, and included the sum in the SED fit. We deferred further analysis to follow-up work with forthcoming deeper Spitzer–IRAC data and follow-up optical and near-infrared photometry.

We used Code Investigating GALaxy Emission (CIGALE)[45,46] for SED fitting of the combined photometry of the source. The SED modelling assumes a single-component star-formation history and solar metallicity[47]. A Chabrier[48] initial mass function is assumed. The resulting best-fitting SED is shown in Extended Data Fig. 6. The infrared luminosity (at 8–1,100 μm) is $(8.0 \pm 1.0) \times 10^{13} L_{\odot}$ (where $L_{\odot}$ is the luminosity of the Sun).

Protocluster comparison sample. To place SPT2349-56 in context and to compare it with other systems claimed to be protoclusters, we identified from the literature various SMG-rich overdensities at redshifts between 2 and 5. Although a direct comparison of the number counts (number per deg^2) of SMG-overdense systems can be performed, it involves making somewhat arbitrary choices of enclosed areas and redshift boundaries. We opted in Fig. 2 to show instead a curve of growth analysis of the 870-μm flux density. We considered only galaxies that have been confirmed to be protocluster members via spectroscopic redshifts. The data were drawn from a recent compilation[49] and original references therein.

The Great Observatories Origins Deep Survey North (GOODS-N) overdensity at $z = 1.99$ (refs. [9,50,51]) spans a roughly 10′ by 10′ field in the Hubble Deep Field North, and contains nine SMGs within a redshift range (Δz) of 0.008. The probability of this large an overdensity being drawn from the field distribution by chance is less than 0.01%. Interestingly, only a modest overdensity of Lyman-break galaxies is found in this GOODS-N structure.

The Cosmic Evolution Survey (COSMOS) $z = 2.5$ SMG overdensity[8] is similar to the GOODS-N structure in terms of the numbers and luminosities of the component SMGs, the angular size of the system, and the modest overdensity of LBGs associated with it.

The overdensity in the survey MRC1138 was originally discovered as an overdensity of Lyα and Hα emitters[52]. Follow-up observations[53,54] revealed the presence of five SMGs, in addition to an active galactic nucleus (AGN) known as the Spiderweb Galaxy. This is a radio-loud AGN that resides in a large Lyα halo.

The SSA22 protocluster was one of the first discovered by observing an overdensity of Lyman-break galaxies[55]. It is an extremely extended structure located at $z = 3.09$, with Lyα emitters spanning more than 50 cMpc (ref. [56]). Submillimetre observations of the field have revealed a population of at least eight SMGs[10,50,57–59].

The COSMOS $z = 2.1$ protocluster[60] lacks enough deep 850-μm data to characterize the Herschel–SPIRE sources identified in the structure. We estimated 870-μm flux densities by taking the published infrared luminosities (L_{IR}; integrated over 3–1,100 μm) and used the SED of the galaxy collision Arp 220 to estimate $S_{870μm}$, finding that $L_{IR} = 2 \times 10^{12} L_{\odot}$ corresponds to $S_{870μm} = 1$ mJy at $z \approx 2$. For the SSA22 protocluster, we used the measured 870-μm flux density when available and otherwise estimate it from the 1.1-mm flux using a standard conversion at $z \sim 3$ of $S_{870μm} = 2 \times S_{1.1mm}$. To create the curves of growth for Fig. 2, we defined the centre of each protocluster by computing the median right ascension and declination of all submillimetre sources. We checked that randomly adjusting the centres of the curve-of-growth tracks by about 1′ did not boost the curves by more than 10%, showing that the curves of growth for the literature SMG overdensities are insensitive to the adopted centre.

Recently, there have also been detections of SMG overdensities at $z \approx 4$. The first, GN20, at $z = 4.05$, was discovered through the serendipitous detection of CO(4–3) emission from two SMGs[61], with two further SMGs detected subsequently[62]. An excess of *B*-band dropouts was also observed in this structure, several of these dropouts being confirmed spectroscopically to lie at $z \approx 4.05$. The second SMG overdensity at a redshift greater then 4, HDF850.1, contains a single SMG, a quasar and 11 spectroscopically confirmed galaxies. This SMG has a confirmed redshift of $z = 5.18$ (ref. [63]). The AzTEC-3 overdensity is centred on a single SMG at $z = 5.3$,

with 12 spectroscopically confirmed optical galaxies at the same redshift. This is a relatively dense structure, with most of the galaxies residing within a circle of about 1′ in diameter. The most luminous example at $z \approx 4$, SMM J004224, was found from the Herschel surveys[23], with several additional 870-μm sources in the surrounding field. In Fig. 2 we also plot the flux from the satellites that have observed SPIRE colours consistent with $z = 4$ (ref. [64]).

Overdensities of SMGs and optical galaxies have also been found around high-redshift radio galaxies (HzRGs)[65], continuing to confirm HzRGs as useful beacons of structure forming in the early Universe. However, none of these systems comes close to the level of overdensity found in SPT2349-56, and they also suffer from the bias inherent in targeting these sources—namely, that one or more protocluster members has to be radioluminous.

There have also been discoveries of compact binary hyperluminous infrared galaxy (HyLIRG) systems, the most luminous of which is the $z = 2.4$ source HATLAS J084933 (ref. [66]), with others approaching this luminosity[67,68]. In each of these systems, the dynamics and SFRs are dominated by two SMGs, but there is no strong evidence of any surrounding protocluster in the form of an excess of galaxies selected optically or in the submillimetre. In one case[68], there is evidence for a relative void around the structure. These systems may simply be instances of very rare events in fairly typical (but still massive) halos[69], analogous to hyperluminous quasars[70].

Theoretical studies of N-body simulations have shown that the progenitors of $z = 0$ dark-matter halos with masses greater than $10^{15} M_\odot$ should extend to effective radii greater than about 5 cMpc at redshifts higher than 2 (refs. [12,26]). Given that the overdensities listed above are mostly concentrated in small areas, it is difficult to asses their exact evolution or to compare them easily with simulated structures. Interpreting a small overdense region at high redshift as a 'protocluster core' is certainly prone to misinterpretation, and small overdense regions at high redshift can evolve into halos spanning a range of masses at the current epoch[12]. Chiang et al.[12] suggest investigating if an overdensity extends to larger scales (more than 20 cMpc), in order to better determine whether it will form a cluster with a mass greater than $10^{15} M_\odot$. But this is difficult at high redshift, because the excess of galaxies will be less pronounced on larger scales, and it is challenging to detect high-redshift, low-luminosity galaxies.

To assess the evolution of SPT2349-56, we also compared it with a selection of the highest-mass galaxy clusters at lower redshifts (less than around 2.5). The clusters XMMU J2235.3 (ref. [71]) and XLSSC 122 (ref. [2]) were both discovered in X-ray surveys with masses derived from the X-ray light profile. Meanwhile IDCS J1426.5 + 3508 (refs. [3,72]), JKCS5041 (ref. [73]), CL J1449 + 0856 (refs. [74,75]) and CL J1001 + 0220 (ref. [1]) were discovered as overdensities of massive red galaxies and have corresponding X-ray detections. Masses for all were estimated using these X-ray detections and are consistent with dynamical masses estimated using the velocity dispersion. We also show the sample of clusters discovered in the 2,500 deg² SPT survey using the Sunyaev–Zel'dovich effect, with mass estimates based on the velocity dispersions[76].

Geometry and dynamics. In Extended Data Fig. 7 we investigate the geometry and kinematics of the SPT2349-56 system. By analysing the velocity distribution relative to the mean redshift of 4.304, ΔV, versus projected radius for our sample of 14 galaxies, we can see that—even with our most conservative mass estimate—at least 12 of the 14 galaxies appear to be bound. We show the escape velocity as a function of radius for a point source and a Navarro–Frenk–White (NFW) profile with a mass of $1.16 \times 10^{13} M_\odot$. The NFW profile assumes a virial radius of 200 kpc and a concentration of 5—values typical of massive halos at this epoch, found in the N-body simulations described in the section on 'Simulations' below. Even if the projected radius were to account for only a fraction of the true physical distance (that is, separated along the line of sight), it is likely that most of the galaxies with low relative velocity (ΔV less than $500 \, \mathrm{km \, s^{-1}}$) will still be bound and eventually collapse into a single galaxy. The cumulative distribution of relative velocities is also shown, alongside that of a Gaussian distribution with $\sigma = 408 \, \mathrm{km \, s^{-1}}$. The observed distribution is smooth and well fit by the Gaussian. The relative kinematics of the 14 galaxies is consistent with most, if not all, members being mutually bound with SPT2349-56, following a Gaussian ΔV distribution characteristic of a virialized system.

An alternative interpretation of this system is that we are observing individual galaxies, or separate less-massive halos along a filament aligned with our line of sight. The probability of this occurring might be boosted by our selection technique, given that the beam size of SPT is approximately 1′. We are preferentially sensitive to structure on this scale, and require several of the most luminous galaxies found in the Universe together in the SPT beam to exceed our detection threshold.

To explore this possibility, we also show in Extended Data Fig. 7 the most extreme interpretation of SPT2349-56, whereby none of the velocity offsets is peculiar, and the physical distribution of the SPT2349-56 galaxies is represented by the relative velocities being entirely due to cosmic expansion rather than

peculiar motions. In this case, the galaxies of course are maximally separated along a (proper distance) 3-Mpc filament (compared with their 130-kpc maximum tangential extent). However, this cannot be the true distribution, and a more realistic interpretation would entail some 20–30% of velocities being Hubble flow ($600 \times 900 \, \mathrm{pc}$ extent), and the remainder representing peculiar motions. (Note that this would be true for any massive halo at this epoch, given the expected velocity dispersion of about $400 \, \mathrm{km \, s^{-1}}$ and Hubble constant of $470 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$.) This explanation, if substantiated, would alter our interpretation of the system, specifically decreasing the mass estimate as the velocity offsets are not due to peculiar motions and the system is not virialized.

However, given the observed kinematics and spatial configuration of the galaxies discussed above, we argue that SPT2349-56 is less likely to consist of multiple groups, widely separated on a line-of-sight filament. To back up this claim, we performed a systemic search for similar filamentary structures in the N-body simulations described in 'Simulations' below. We first searched for groups at redshifts greater than 4 with more than eight galaxies within a projected radius of more than 150 kpc but can be extended up to 5 Mpc along the line of sight. There are three such systems in our 1-cGpc³ simulation, and they are also displayed in Extended Data Fig. 7, showing their geometry. Each of these three analogous systems in the simulation has a total halo mass of about $10^{13} M_\odot$. None is extended much past 500 kpc, suggesting that, at redshifts of around 4, there are no substantially extended filaments hosting massive galaxies in the simulation that are comparable to even the gas masses of our SPT2349-56 ALMA galaxies. These simulated structures are not filaments; they are close to collapsed structures that are slightly cigar-shaped, and although SPT2349-56 could in principle be distributed like this, it does not fundamentally change our discussion here. (We also note that we are plotting two different things in Extended Data Fig. 7: velocity offsets for SPT2349-56, and the actual geometry of the simulated galaxies.) These results further suggest that most of the velocity offsets in the SPT2349-56 galaxies are truly due mostly to peculiar motion rather than line-of-sight projection. At least one of the simulation systems found appears to be a chance projection of multiple groups, but it has a characteristically different ΔV distribution. The system has multiple smaller groups with $\sigma \approx 200 \, \mathrm{km \, s^{-1}}$, reflecting their smaller masses, separated in velocity space by $500 \, \mathrm{km \, s^{-1}}$.

We also note that a filamentary structure feeding a halo of approximate mass $10^{13} M_\odot$ should have a characteristic width of 200 kpc to 400 kpc (refs. [77,78]), compared with the 80 kpc (FWHM) width we find for our 14 ALMA galaxies. This is suggestive that, even with a direct line-of-sight view down a filament, the configuration is far more concentrated than one would infer based on the typical size of filament. We then ask what sort of environment a roughly 80-kpc-wide filament (consistent with our SPT2349-56 galaxies) would connect, and infer a typical halo mass of around $2 \times 10^{12} M_\odot$ (ref. [77]). Given that the SPT2349-56 galaxies contain at least this much mass entirely in their cold gas (as we have adopted the lowest plausible $\alpha_{\mathrm{CO}} = 0.8$ conversion factor), and are probably hosted by haloes at least twice as large as this, it becomes somewhat contradictory that they could be found along a filament that is connecting a much less massive halo.

A further important piece of evidence against these 14 galaxies being a filament is that we find no additional sources in the surrounding spectral windows in either side band, in the band-7 or the band-3 ALMA data (Extended Data Fig. 7d). Although all 14 sources are clustered within $1,500 \, \mathrm{km \, s^{-1}}$ of each other, our full frequency coverage extends over relative velocities of $6,500 \, \mathrm{km \, s^{-1}}$ (for [C II]) and $27,000 \, \mathrm{km \, s^{-1}}$ (for CO(4–3)), or many tens of Mpc line-of-sight distance. If SPT2349-56 were actually an elongated filament, one might expect to see a distribution of sources over a much larger fraction of our observed frequency bandwidth. Although we can not rule out the possibility that we are observing multiple galaxies along our line of sight, the observed geometry and kinematics suggest that SPT2349-56 consists of a single group of mutually gravitationally bound galaxies.

Another way to assess the configuration of this system is by using the Millennium Simulation[4] database, with the implementation of the galaxy-formation recipe[79]. We searched the model output at $z \approx 4$ for galaxies with a total baryonic mass in excess of around $10^{11} M_\odot$, consistent with the combined mass of gas and stars estimated for the brightest ALMA-resolved galaxies in SPT2349-56. We used this total baryonic mass cut as it meant we were less sensitive to the details of the early star-formation histories of galaxies in the model. We found only one galaxy in the $3.2 \times 10^8 \, \mathrm{cMpc^3}$ simulation volume at $z = 4$ with a total baryonic mass above $10^{11} M_\odot$. Half of the baryonic mass in this model galaxy is found in stars, and it has a $3 \times 10^8 M_\odot$ black hole. Moreover, this galaxy is the central galaxy of a $1 \times 10^{13} M_\odot$ halo—the optimal environment for finding merging galaxies, according to simulations[80]. Searching the environment of this system, we found another four massive galaxies are distributed across a region with a roughly 0.5-cMpc diameter around the central galaxy, with baryonic masses of more than $7 \times 10^{10} M_\odot$ (more than 5% of the primary). By redshift zero, these galaxies are all predicted to reside in a $1 \times 10^{15} M_\odot$ halo, consistent with our other assessments of the outcome of this system.

Simulations. To further place SPT2349-56 in context, we compared it with the predictions of a theoretical model for SMG overdensities[69,81]. We used the MultiDark[82] *N*-body simulation, which is one of the largest (2.91 Gpc³) available *N*-body simulations that still resolves SMG-like halos (with halo masses of more than roughly $10^{12} M_\odot$). We analysed the $z = 4.68$ and $z = 4.1$ snapshots, which are the available snapshots that are closest in redshift to SPT2349-56. We created halo catalogues using the Rockstar halo finder[83], and assigned stellar masses to dark-matter halos using a relation derived based on a sub-halo abundance-matching relation[84]. To assign SFRs, it is assumed that the distribution of specific SFRs (SFRs per unit stellar mass, hereafter SSFR) is the sum of two Gaussians, corresponding to quiescently star-forming and starburst galaxies[85]. The median SSFR value is based on the abundance-matching-derived relation[84], and the starburst fraction and the widths of the Gaussian distributions are set on the basis of observations of massive, high-redshift star-forming galaxies similar to the members of SPT2349-56 (ref. [85]). The dust mass, M_d, is estimated from the stellar mass by using empirical gas-fraction and metallicity relations[86]. Once SFR, stellar mass (M_*), and M_d values are assigned to each halo, $S_{870\mu m}$ is calculated using the following fitting function, which was derived from results obtained by performing dust radiative transfer calculations on hydrodynamical simulations of both isolated and interacting galaxies[81,87]:

$$ S_{870mm} = 0.81 \text{ mJy} \left(\frac{\text{SFR}}{100\ M_\odot \text{yr}^{-1}} \right)^{0.43} \left(\frac{M_d}{10^8 M_\odot} \right)^{0.54} $$

where $S_{870\mu m}$ is the 870-μm flux density, SFR is the star-formation rate, and M_d is the dust mass. A scatter of 0.13 dex is included when applying the relation.

Once $S_{870\mu m}$ was assigned to each halo, we searched the entire simulation volume for the most luminous regions. We began at each independent halo and calculated the total $S_{870\mu m}$ of all halos within a given projected radius, r, and line-of-sigh distance along a given axis of this halo. We used a line-of-sight distance of 1 Mpc for the $z = 4$ snapshot and 2 Mpc for the $z = 2.5$ snapshot, reflecting the Hubble constant at each epoch and the expected velocity dispersion of roughly 400 km s⁻¹. For each value of r, we recorded the largest total $S_{870\mu m}$ obtained (across all halos). We performed 100 Monte Carlo iterations for each snapshot; in each iteration, galaxy properties were re-assigned, drawing from the distributions described above. The shaded region in Fig. 2 shows the entire region spanned by the 100 realizations of the maximum $S_{870\mu m}$ versus area curves. To compare SPT2349-56 with lower-redshift protoclusters, we performed a similar analysis on a snapshot at $z = 2.49$ with 20 Monte Carlo iterations.

Data availability. Data can be made available upon reasonable request to T.B.M.

27. Carlstrom, J. E. et al. The 10 meter South Pole Telescope. *Publ. Astron. Soc. Pacific* **123**, 568 (2011).
28. Ashby, M. L. N. et al. The Spitzer South Pole Telescope deep field: survey dDesign and infrared array camera catalogs. *Astrophys. J. Suppl. Ser.* **209**, 22 (2013).
29. Planck Collaboration. Planck 2013 results. XXII. Constraints on inflation. *Astron. Astrophys.* **571**, A22 (2014).
30. Weiß, A. et al. ALMA redshifts of millimeter-selected galaxies from the SPT survey: the redshift distribution of dusty star-forming galaxies. *Astrophys. J.* **767**, 88 (2013).
31. Siringo, G. et al. The Large APEX BOlometer CAmera LABOCA. *Astron. Astrophys.* **497**, 945–962 (2009).
32. Weiß, A. et al. LABOCA observations of nearby, active galaxies. *Astron. Astrophys.* **490**, 77–86 (2008).
33. Aravena, M. et al. The ALMA spectroscopic survey in the Hubble ultra deep field: search for CII line and dust emission in $6 < z < 8$ galaxies. *Astrophys. J.* **833**, 71 (2016).
34. Williams, J. P. et al. Clumpfind: determining structure in molecular clouds (Astrophysics Source Code Library ascl:1107.014, 2011).
35. Wilson, W. E. et al. The Australia Telescope Compact Array broad-band backend: description and first results. *Mon. Not. R. Astron. Soc.* **416**, 832–856 (2011).
36. Sault, R. J., Teuben, P. J. & Wright, M. C. H. A retrospective view of Miriad. *Astron. Data Analysis Software Systems IV* **77**, 433 (1995).
37. Mcmullin, J. P., Waters, B., Schiebel, D., Young, W. & Golap, K. CASA architecture and applications. *Astron. Data Analysis Software Systems XVI* **376**, 127 (2007).
38. Solomon, P. M. et al. The molecular interstellar medium in ultraluminous infrared galaxies. *Astrophys. J.* **478**, 144–161 (1997).
39. Fazio, G. G. et al. The Infrared Array Camera (IRAC) for the Spitzer Space Telescope. *Astrophys. J. Suppl. Ser.* **154**, 10–17 (2004).
40. Werner, M. W. et al. The Spitzer Space Telescope Mission. *Astrophys. J. Suppl. Ser.* **154**, 1–9 (2004).
41. Béthermin, M., Dole, H., Cousin, M. & Bavouzet, N. Submillimeter number counts at 250, 350 and 500 microns in BLAST data. *Astron. Astrophys.* **516**, A43 (2010).
42. Barger, A. J. et al. Is there a maximum star formation rate in high-redshift galaxies? *Astrophys. J.* **784**, 9 (2014).
43. Safarzadeh, M., Hayward, C. C., Ferguson, H. C. & Somerville, R. S. What shapes the far-infrared spectral energy distributions of galaxies? *Astrophys. J.* **818**, 62 (2016).
44. Bothwell, M. S. et al. A survey of molecular gas in luminous sub-millimetre galaxies. *Mon. Not. R. Astron. Soc.* **429**, 3047–3067 (2013).
45. Burgarella, D. et al. Star formation and dust attenuation properties in galaxies from a statistical ultraviolet-to-far-infrared analysis. *Mon. Not. R. Astron. Soc.* **360**, 1413–1425 (2005).
46. Noll, S. et al. Analysis of galaxy spectral energy distributions from far-UV to far-IR with CIGALE: studying a SINGS test sample. *Astron. Astrophys.* **507**, 1793–1813 (2009).
47. Bruzual, G. & Charlot, S. Stellar population synthesis at the resolution of 2003. *Mon. Not. R. Astron. Soc.* **344**, 1000–1028 (2003).
48. Chabrier, G. The galactic disk mass function: reconciliation of the Hubble Space Telescope and nearby determinations. *Astrophys. J.* **586**, L133–L136 (2003).
49. Casey, C. M. The ubiquity of coeval starbursts in massive galaxy cluster progenitors. *Astrophys. J.* **824**, 36 (2016).
50. Chapman, S. C., Blain, A. W., Smail, I. & Ivison, R. J. A redshift survey of the submillimeter galaxy population. *Astrophys. J.* **622**, 772–796 (2005).
51. Blain, A. W., Chapman, S. C., Smail, I. & Ivison, R. Clustering of submillimetre-selected galaxies. *Astrophys. J.* **611**, 725–731 (2004).
52. Kurk, J. D. et al. A search for clusters at high redshift. I. Candidate Lyalpha emitters near 1138-262 at $z = 2.2$. *Astron. Astrophys.* **358**, L1–L4 (2000).
53. Kuiper, E. et al. A SINFONI view of flies in the Spiderweb: a galaxy cluster in the making. *Mon. Not. R. Astron. Soc.* **415**, 2245–2256 (2011).
54. Dannerbauer, H. et al. An excess of dusty starbursts related to the Spiderweb galaxy. *Astron. Astrophys.* **570**, A55 (2014).
55. Steidel, C. et al. A large structure of galaxies at redshift $z \sim 3$ and its cosmological implications. *Astrophys. J.* **492**, 428–438 (1998).
56. Hayashino, T. et al. Large-scale structure of emission-line galaxies at $z = 3.1$. *Astron. J.* **128**, 2073–2079 (2004).
57. Chapman, S. C. et al. Submillimeter imaging of a protocluster region at $z = 3.09$. *Astrophys. J.* **548**, L17–L21 (2001).
58. Geach, J. E. et al. A submillimetre survey of Lyman α haloes in the SA 22 protocluster at $z = 3.1$. *Mon. Not. R. Astron. Soc.* **363**, 1398–1408 (2005).
59. Umehata, H. et al. ALMA deep field in SSA22: a concentration of dusty starbursts in a $z = 3.09$ protocluster core. *Astrophys. J. Lett.* **815**, L8 (2015).
60. Hung, C.-L. et al. Large scale structure around a $z = 2.1$ cluster. *Astrophys. J.* **826**, 130 (2016).
61. Daddi, E. et al. Two bright submillimeter galaxies in a $z = 4.05$ protocluster in GOODS-North, and accurate radio-infrared photometric redshifts. *Astrophys. J.* **694**, 1517–1538 (2009).
62. Daddi, E. et al. A CO emission line from the optical and near-IR undetected submillimeter galaxy GN10. *Astrophys. J.* **695**, L176–L180 (2009).
63. Walter, F. et al. The intense starburst HDF 850.1 in a galaxy overdensity at $z \sim 5.2$ in the Hubble Deep Field. *Nature* **486**, 233–236 (2012).
64. Lewis, A. J. R. et al. Ultra-red galaxies signpost candidate proto-clusters at high redshift. Preprint at https://arxiv.org/pdf/1711.08803v1.pdf (2017).
65. Noirot, G. et al. HST grism confirmation of two $z \sim 2$ structures from the clusters around radio-loud AGN (CARLA) survey. *Astrophys. J.* **830**, 90 (2016).
66. Ivison, R. J. et al. Herschel-ATLAS: a binary HyLIRG pinpointing a cluster of starbursting proto-ellipticals. *Astrophys. J.* **772**, 137 (2013).
67. Fu, H. et al. The rapid assembly of an elliptical galaxy of 400 billion solar masses at a redshift of 2.3. *Nature* **498**, 338–341 (2013).
68. Chapman, S. C. et al. A millimetre-wave redshift search for the unlensed HyLIRG, HS1700.850.1. *Mon. Not. R. Astron. Soc.* **453**, 951–959 (2015).
69. Miller, T. B., Hayward, C. C., Chapman, S. C. & Behroozi, P. S. The bias of the submillimetre galaxy population: SMGs are poor tracers of the most-massive structures in the $z \sim 2$ Universe. *Mon. Not. R. Astron. Soc.* **452**, 878–883 (2015).
70. Trainor, R. F. & Steidel, C. C. The halo masses and galaxy environments of hyperluminous QSO's at $z \sim 2.7$ in the Keck Baryonic Structure Survey. *Astrophys. J.* **752**, 39 (2012).
71. Rosati, P. et al. Multi-wavelength study of XMMU J2235.3-2557: the most massive galaxy cluster at $z > 1$. *Astron. Astrophys.* **508**, 583–591 (2009).
72. Brodwin, M. et al. IDCS J1426.5+3508: the most massive galaxy cluster at $z > 1.5$. *Astrophys. J.* **817**, 122 (2016).
73. Andreon, S. et al. JKCS 041: a Coma cluster progenitor at $z = 1.803$. *Astron. Astrophys.* **565**, A120 (2014).
74. Gobat, R. et al. A mature cluster with X-ray emission at $z = 2.07$. *Astron. Astrophys.* **526**, A133 (2011).
75. Gobat, R. et al. WFC3 GRISM confirmation of the distant cluster Cl J1449+0856 at $< z > = 2.00$: quiescent and star-forming galaxy populations. *Astrophys. J.* **776**, 9 (2013).
76. Bocquet, S. et al. Mass calibration and cosmological analysis of the SPT-SZ galaxy cluster sample using velocity dispersion and X-ray measurements. *Astrophys. J.* **799**, 214 (2014).
77. Mandelker, N., van Dokkum, P. G., Brodie, J. P., van den Bosch, F. C. & Ceverino, D. Cold filamentary accretion and the formation of metal poor globular clusters and halo stars. Preprint at https://arxiv.org/abs/1711.09108v1 (2017).
78. Cautun, M., van de Weygaert, R., Jones, B. J. T. & Frenk, C. S. Evolution of the cosmic web. *Mon. Not. R. Astron. Soc.* **441**, 2923–2973 (2014).
79. Bower, R. G. et al. Breaking the hierarchy of galaxy formation. *Mon. Not. R. Astron. Soc.* **370**, 645–655 (2006).
80. Hopkins, P. F., Hernquist, L., Cox, T. J. & Kereš, D. A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. I. Galaxy mergers and quasar activity. *Astrophy. J. Suppl. Ser.* **175**, 356–389 (2008).

81. Hayward, C. C. et al. Spatially unassociated galaxies contribute significantly to the blended submillimetre galaxy population: predictions for follow-up observations of ALMA sources. *Mon. Not. R. Astron. Soc.* **434**, 2572–2581 (2013).

82. Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E. & Primack, J. Halo concentrations in the standard Λ cold dark matter cosmology. *Mon. Not. R. Astron. Soc.* **423**, 3018–3030 (2012).

83. Behroozi, P. S., Wechsler, R. H. & Wu, H.-Y. The rockstar phase-space temporal halo finder and the velocity offsets of cluster cores. *Astrophys. J.* **762**, 109 (2011).

84. Behroozi, P. S., Wechsler, R. H. & Conroy, C. The average star formation histories of galaxies in dark matter halos from $z = 0-8$. *Astrophys. J.* **770**, 57 (2013).

85. Sargent, M. T. et al. The contribution of starbursts and normal galaxies to infrared luminosity functions at $z < 2$. *Astrophys. J.* **747**, L31 (2012).

86. Hayward, C. C. et al. Submillimetre galaxies in a hierarchical universe: number counts, redshift distribution and implications for the IMF. *Mon. Not. R. Astron. Soc.* **428**, 2529–2547 (2013).

87. Hayward, C. C. et al. What does a submillimeter galaxy selection actually select? The dependence of submillimeter flux density on star formation rate and dust mass. *Astrophys. J.* **743**, 159 (2011).
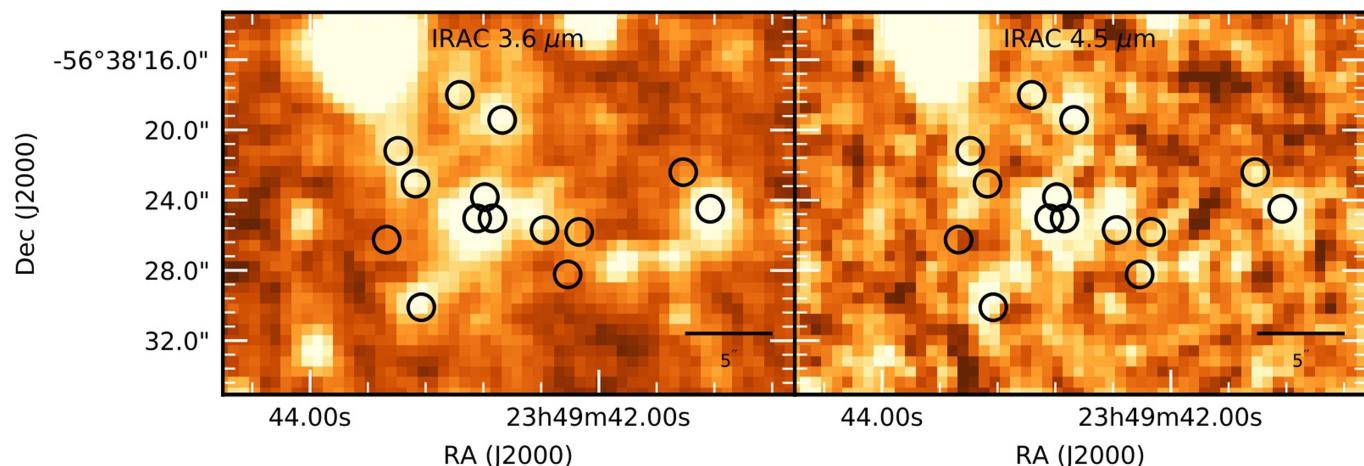
Extended Data Fig. 1 | IRAC observations of SPT2349-56. Circles show the locations of the 14 sources detected in ALMA band 7 (see Methods). Nine of the 14 ALMA sources are detected in the IRAC bands with at least 3σ confidence, including the two faintest [C II] sources from the blind line survey.
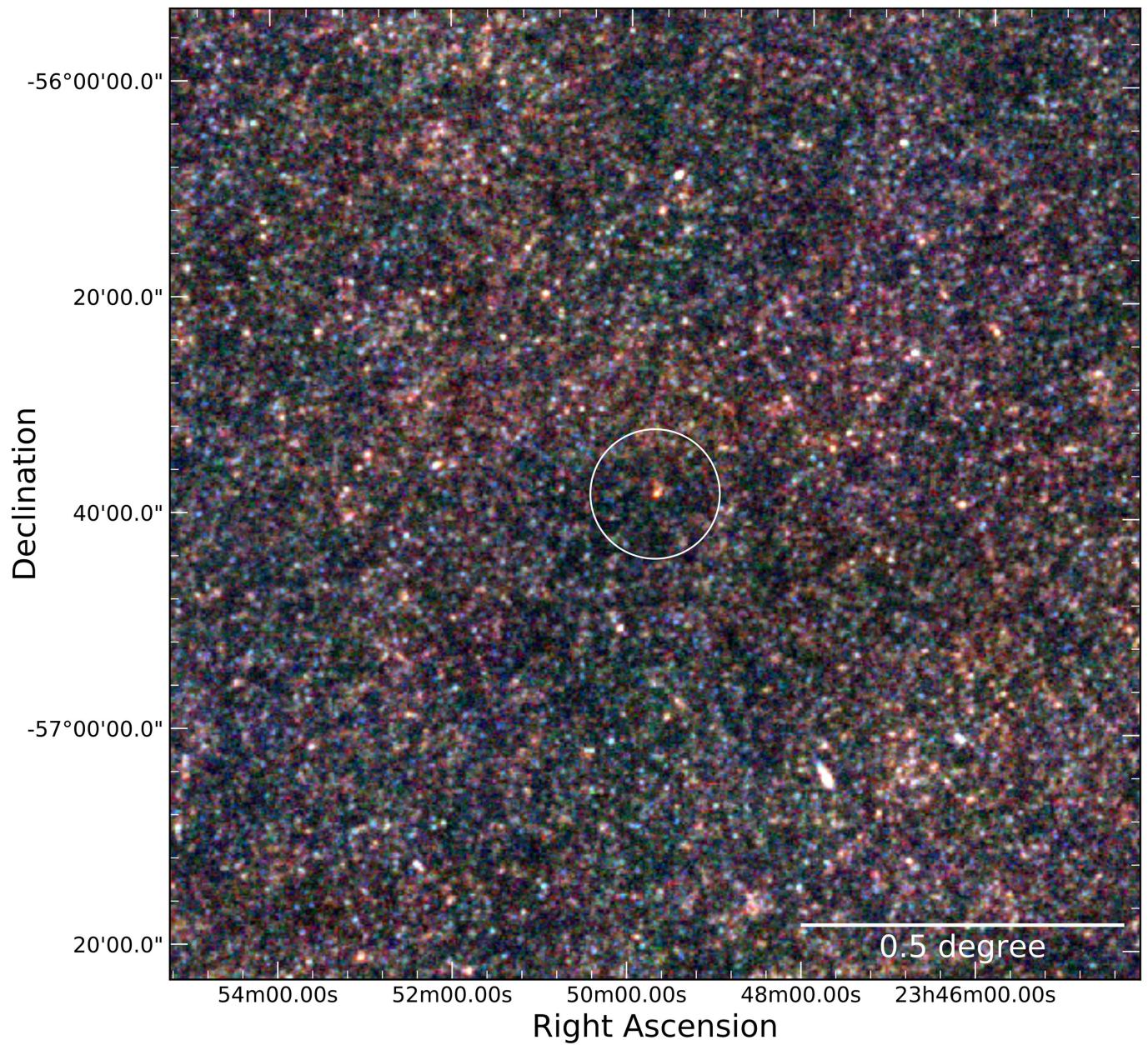
Extended Data Fig. 2 | Herschel–SPIRE image. An RGB scale is used to represent sources selected at wavelengths of 500 μm (red), 350 μm (green) and 250 μm (blue), with the red SPT2349-56 extended complex clearly visible in a relative void in the foreground $z \approx 1$ cosmic infrared background (blue to green coloured galaxies).
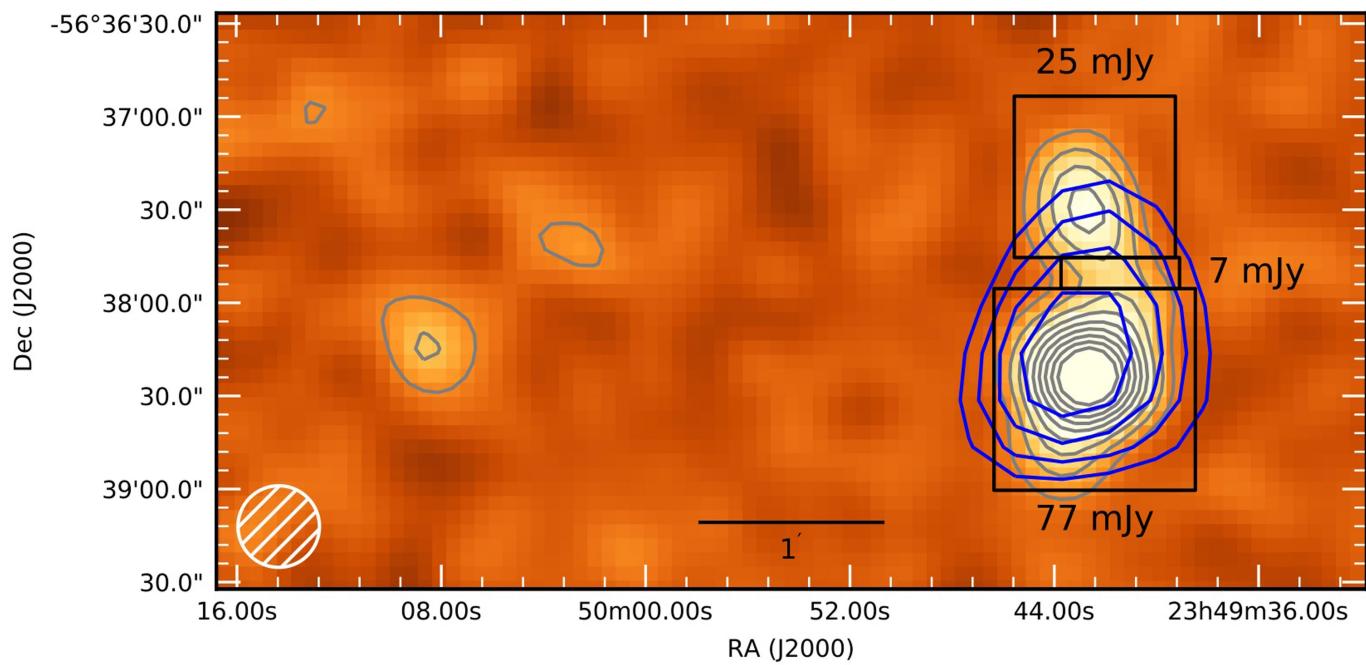
Extended Data Fig. 3 | Wide-field 870-μm image and photometry.
A wide-field LABOCA image (21″ beam size; white circle) of SPT2349-56. The image r.m.s. noise is 1.3 mJy at the centre of region shown to the right, increasing to 2 mJy at the edges of this region. The total flux density recovered is 110.0 ± 9.5 mJy. Subregions are drawn with black outlines, showing three different regions and their recovered flux densities. Grey contours start at 3σ and increase in steps of 3σ. SPT 1.4-mm contours are also shown (blue), revealing that even with the 1′ beam of SPT, SPT2349-56 is resolved. One additional submillimetre source is detected at $> 5\sigma$ in the LABOCA image to the east (left) of the primary source, though Herschel–SPIRE photometry indicates that it is unlikely to be at $z \approx 4$.
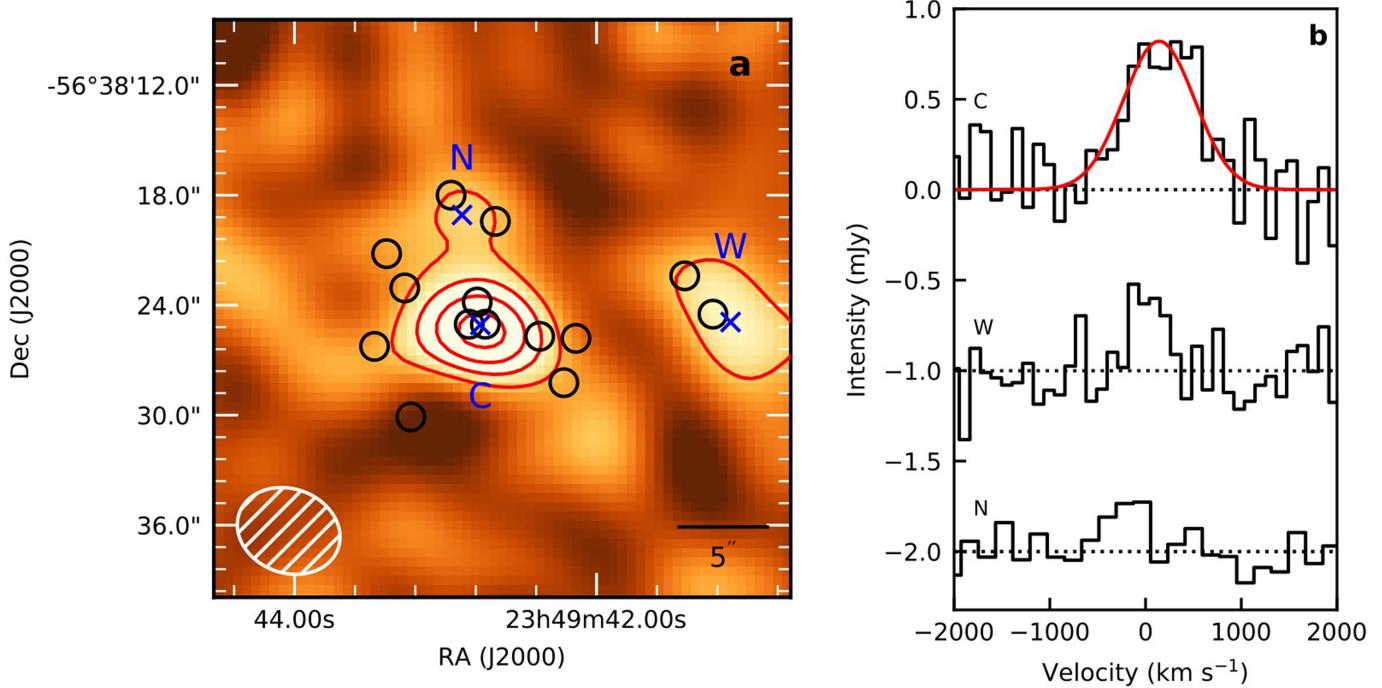
Extended Data Fig. 4 | CO(2–1) observations of SPT2349-56. a, The colour map and red contours trace the CO(2–1) line integrated over the central 830 km s^{-1}, with the contours spaced by 2σ, starting at 2σ. The grey contours show the 1.1-mm ALMA continuum detections. N, north; W, west; C, centre. Black circles show the location of the 14 [C II]/continuum sources identified with ALMA. **b**, One-dimensional spectra extracted at the positions indicated with blue crosses in panel **a**.
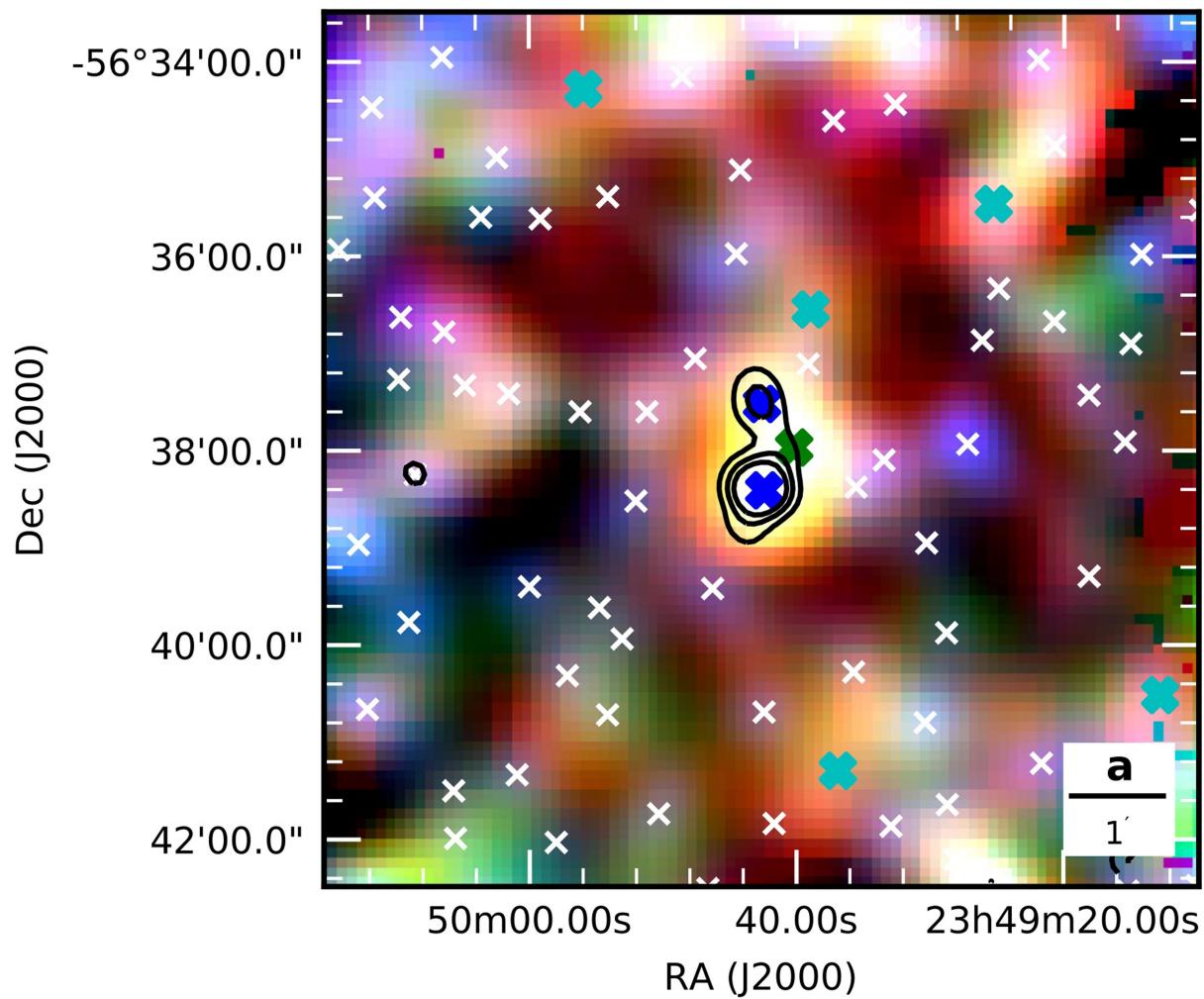
Extended Data Fig. 5 | SPIRE RGB image and source colours in the field surrounding SPT2349-56. a, A deep SPIRE false-colour image is shown with LABOCA contours overlaid. Locations of the 250-μm peaks used for analysis are marked with crosses (the faintest are not visible because of the contrast adopted in the image). The turquoise, blue, green and white crosses relate to the turquoise, blue, green and black lines used in panels **b** and **c. b, c**, Colour–colour (**b**) and colour–flux (**c**) diagrams for the 250-μm sources. Error bars represent 1σ standard deviation. The colour–colour diagram shows sources with SNR(250 μm) ratios of

more than 3 and is dominated by the $z \approx 1$ cosmic infrared background in the foreground of SPT2349-56 (sources with colours ranging from blue to green). The colour–flux diagram applies an additional cut for SNR(500 μm) ratios of more than 3. The colour–colour and colour–flux diagrams show that one of three peaks associated with SPT2349-56 is probably a lower-redshift interloper (green symbol), but also that there are five additional sources (blue symbols) in the surrounding region with colours ($S_{500\mu m} > S_{350\mu m} > S_{250\mu m}$) that are suggestive of $z = 4.3$.
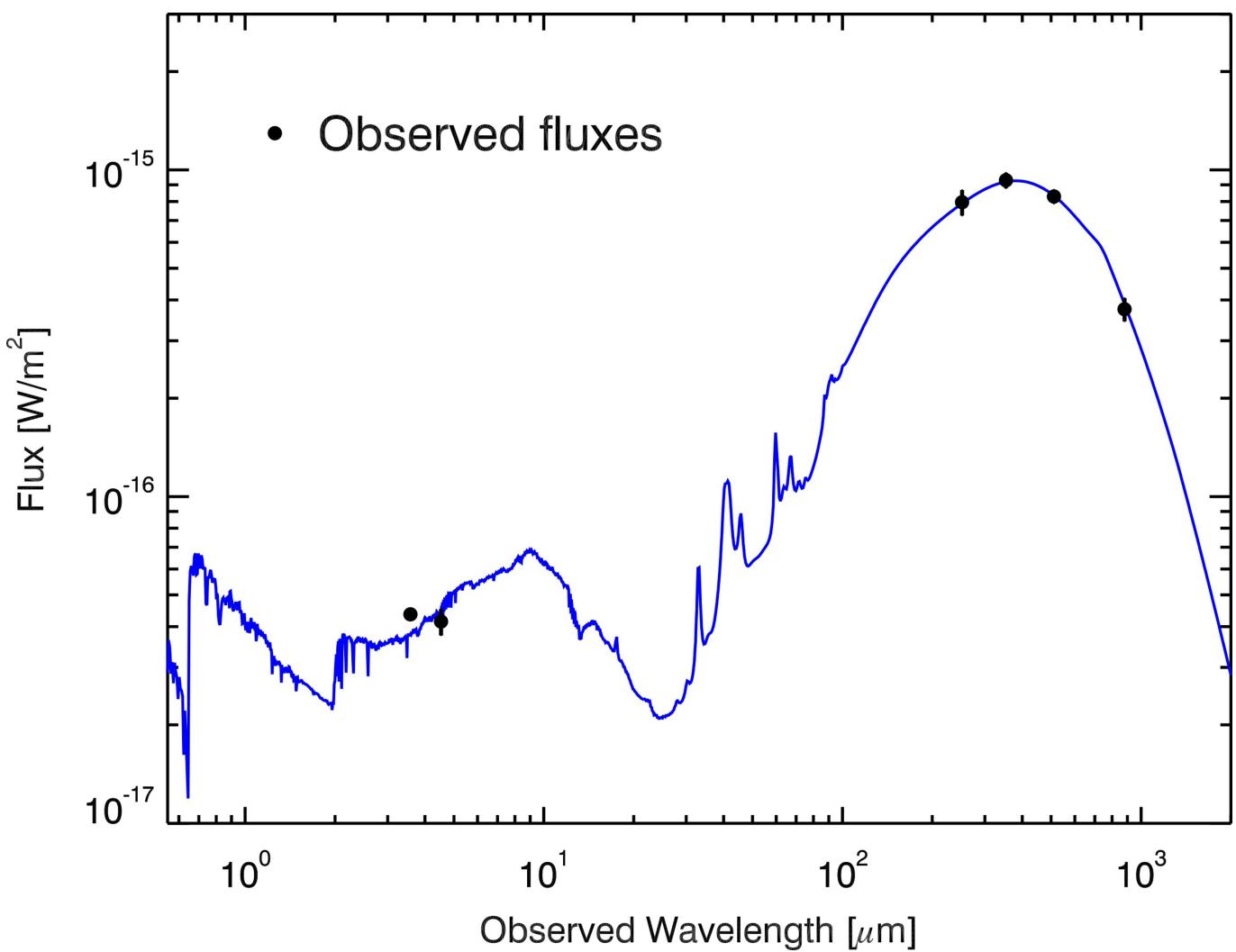
Extended Data Fig. 6 | Spectral energy distribution of SPT2349-56. The SED of the extended SPT2349-56 source is shown, including the summed deconvolved Herschel–SPIRE flux densities, the total 870-μm LABOCA flux density, and the summed IRAC flux densities. Error bars represent 1σ measurement errors. We do not include the SPT 1.4-mm, 2.0-mm and 3.0-mm points because the source is elongated and flux measurements are difficult with the filtering used to make the map. Fitting the SED yields an infrared luminosity of $(8.0 \pm 1.0) \times 10^{13} L_\odot$.
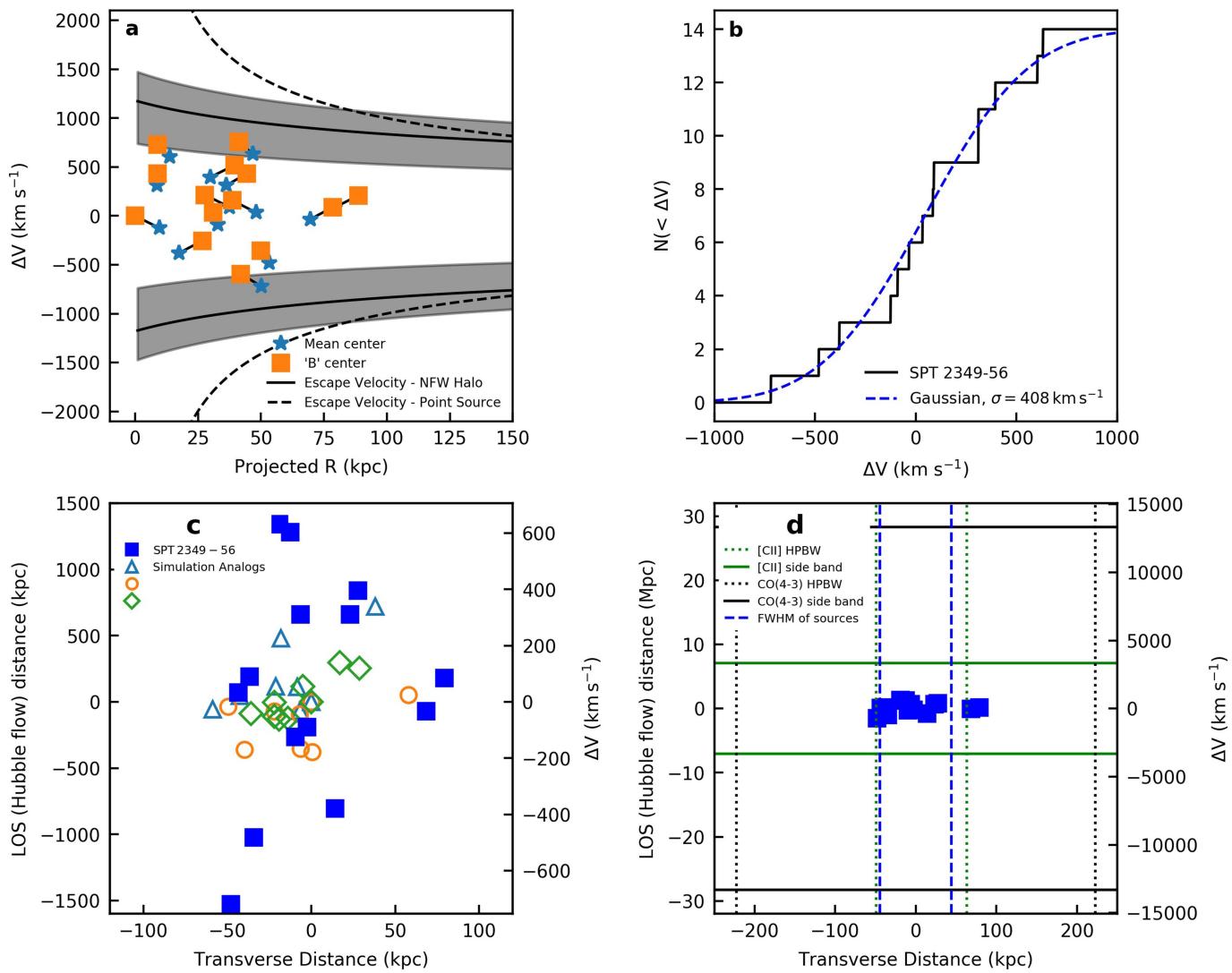
Extended Data Fig. 7 | Geometry and dynamics of the SPT2349-56 system. a, Velocity offsets of the 14 sources versus projected (physical) distance, compared with the escape velocity of a $1.16 \times 10^{13} M_\odot$ NFW halo (virial radius ≈ 200 kpc; concentration $= 5$). The grey shaded region shows our estimated halo-mass uncertainty; also shown is the escape velocity assuming a point mass halo of this same mass. Centres for the distribution of the 14 galaxies are shown at the mean of the distribution, and centred on the 'B' galaxy. All galaxies are bound for all but the lowest range of NFW halo masses. **b**, The cumulative velocity distribution of the SPT2349-56 galaxies, compared with a Gaussian distribution with our estimated dispersion (σ, 408 km s^{-1}), is at least consistent with expectations for a relatively bound system. **c**, The physical distribution of the SPT2349-56 galaxies (blue squares), assuming that their redshifts are due to cosmic expansion rather than peculiar motions. This gives an extreme (but unlikely) possibility that the SPT2349-56 galaxies are stretched out along a filament compared with their 130-kpc maximum tangential extent, but this requires that none of the velocity offsets is a peculiar motion. The open symbols show analogues of SPT2349-56 found when we searched specifically for maximally extended filamentary structures in our N-body simulations. These simulated structures are not filaments; they are instead rather like collapsed structures that are slightly cigar-shaped. The SPT2349-56 galaxies could in principle be distributed like this, but it does not fundamentally change our discussion here. We also note that we are plotting two different things here: velocity offsets for SPT2349-56, and actual geometry (three-dimensional positions) for the simulation galaxies. Even though our search allowed for structures extending by about 5 Mpc along the line of sight (LOS), we found none that stretches beyond 1 Mpc. **d**, As for panel **c**, except that the full extent of our ALMA band 3 and 7 observations is shown. No structures are observed in the sidebands surrounding the 14 observed sources.

Extended Data Table 1 | Observed properties of SPT2349-56 protocluster members

Source	RA (J2000) [h:m:s]	Dec (J2000) [d:m:s]	$S_{1090\,\mu\mathrm{m}}$ [mJy]	$S_{870\,\mu\mathrm{m}}$ [mJy]	CO(4-3) $\int$ S dv [Jy km s^{-1}]	CO(4-3) σ_V [km s^{-1}]	[CII] $\int$ S dv [Jy km s^{-1}]	[CII] σ_V [km s^{-1}]
A	23:49:42.67	−56:38:19.3	4.63 ± 0.04	7.8 ± 0.1	0.99 ± 0.03	376 ± 46	8.81 ± 0.26	354 ± 30
B	23:49:42.79	−56:38:24.0	4.35 ± 0.04	8.2 ± 0.1	0.92 ± 0.03	341 ± 38	7.53 ± 0.22	314 ± 28
C	23:49:42.84	−56:38:25.1	2.69 ± 0.04	6.0 ± 0.1	0.55 ± 0.02	154 ± 13	4.43 ± 0.17	160 ± 10
D	23:49:41.42	−56:38:22.6	2.20 ± 0.08	3.5 ± 0.3	0.69 ± 0.04	485 ± 64	3.62 ± 0.78	346 ± 129
E	23:49:41.23	−56:38:24.4	2.12 ± 0.11	3.3 ± 0.4	0.39 ± 0.02	199 ± 23	3.47 ± 1.24	310 ± 137
F	23:49:42.14	−56:38:25.8	1.69 ± 0.05	3.4 ± 0.1	0.28 ± 0.03	396 ± 103	4.28 ± 0.35	353 ± 35
G	23:49:42.74	−56:38:25.1	1.11 ± 0.04	2.7 ± 0.1	-	-	2.45 ± 0.23	305 ± 50
H	23:49:43.46	−56:38:26.2	0.85 ± 0.05	2.1 ± 0.1	-	-	3.63 ± 0.30	236 ± 31
I	23:49:42.22	−56:38:28.3	0.78 ± 0.05	1.8 ± 0.1	0.18 ± 0.03	277 ± 90	3.18 ± 0.32	236 ± 24
J	23:49:43.22	−56:38:30.1	0.61 ± 0.06	1.6 ± 0.2	0.19 ± 0.02	151 ± 38	3.79 ± 0.29	138 ± 15
K	23:49:42.96	−56:38:17.9	0.34 ± 0.04	1.4 ± 0.1	-	-	2.54 ± 0.17	129 ± 12
L	23:49:42.38	−56:38:25.8	0.23 ± 0.04	0.8 ± 0.1	-	-	2.78 ± 0.20	176 ± 20
M	23:49:43.39	−56:38:21.1	0.21 ± 0.05	0.5 ± 0.2	-	-	1.04 ± 0.14	87 ± 23
N	23:49:43.27	−56:38:22.9	0.18 ± 0.04	0.4 ± 0.1	-	-	0.86 ± 0.16	128 ± 26

Columns 6 and 8 show the integrated line intensity for the CO(4–3) and [C II] lines. Columns 7 and 9 show the line width for the CO(4–3) and [C II] lines.

Extended Data Table 2 | Properties of the three ATCA CO(2–1) sources

ATCA source	ALMA ID	$\int S\,dv$ [Jy km s^{-1}]	σ_V [km s^{-1}]	L′(CO 2-1) 10^{11} [K km s^{-1} pc^2]	M_{gas} [10^{11} M$_\odot$]
Central (C)	B, C, G	0.69 ± 0.076	372 ± 47	1.22 ± 0.12	1.33 ± 0.15
West (W)	D, E	0.16 ± 0.04	166 ± 47	0.29 ± 0.07	0.32 ± 0.08
North (N)	A, K	0.085 ± 0.0028	175 ± 68	0.15 ± 0.05	0.16 ± 0.05

Extended Data Table 3 | Observed properties of all red ($S_{500\,\mu m} > S_{350\,\mu m} > S_{250\,\mu m}$) SPIRE sources in the field surrounding SPT2349-56

RA (J2000) [h:m:s]	Dec (J2000) [d:m:s]	$S_{250\,\mu m}$ [mJy]	$S_{350\,\mu m}$ [mJy]	$S_{500\,\mu m}$ [mJy]	$S_{850\,\mu m}$ [mJy]	$d^{\dagger}$ [arcmin]
23:49:42	−56:38:25	45 ± 3	71 ± 3	96 ± 3	77.0 ± 2.9	-
23:49:43	−56:37:31	21 ± 3	37 ± 3	43 ± 3	25.0 ± 2.8	0.9
23:49:25	−56:35:27	23 ± 4	26 ± 4	32 ± 4	2.9 ± 1.7	5.2
23:49:39	−56:36:33	12 ± 3	16 ± 3	23 ± 4	3.9 ± 1.3	2.1
23:49:36	−56:41:17	7 ± 3	14 ± 3	19 ± 3	3.2 ± 1.6	3.2
23:49:55	−56:34:17	6 ± 4	10 ± 3	20 ± 5	4.8 ± 1.8	5.3
23:49:12	−56:40:31	11 ± 5	16 ± 5	22 ± 5	6.8 ± 2.6	7.7

The LABOCA sources corresponding to SPT2349-56 are listed in the first two rows; the red SPIRE sources in the surrounding field follow. All sources listed are highlighted in Extended Data Fig. 5.
†Distance from central SPT2349-56 source.

LETTER

Remote quantum entanglement between two micromechanical oscillators

Ralf Riedinger[1,3], Andreas Wallucks[2,3], Igor Marinković[2,3], Clemens Löschnauer[1], Markus Aspelmeyer[1], Sungkun Hong[1]* & Simon Gröblacher[2]*

Entanglement, an essential feature of quantum theory that allows for inseparable quantum correlations to be shared between distant parties, is a crucial resource for quantum networks[1]. Of particular importance is the ability to distribute entanglement between remote objects that can also serve as quantum memories. This has been previously realized using systems such as warm[2,3] and cold atomic vapours[4,5], individual atoms[6] and ions[7,8], and defects in solid-state systems[9–11]. Practical communication applications require a combination of several advantageous features, such as a particular operating wavelength, high bandwidth and long memory lifetimes. Here we introduce a purely micromachined solid-state platform in the form of chip-based optomechanical resonators made of nanostructured silicon beams. We create and demonstrate entanglement between two micromechanical oscillators across two chips that are separated by 20 centimetres . The entangled quantum state is distributed by an optical field at a designed wavelength near 1,550 nanometres. Therefore, our system can be directly incorporated in a realistic fibre-optic quantum network operating in the conventional optical telecommunication band. Our results are an important step towards the development of large-area quantum networks based on silicon photonics.

In recent years, nanofabricated mechanical oscillators have emerged as a promising platform for quantum information processing. The field of opto- and electromechanics has seen great progress, including ground-state cooling[12,13], quantum interfaces to optical or microwave modes[14,15], mechanical squeezing[16] and single-phonon manipulation[17–20]. Demonstrations of distributed mechanical entanglement, however, have so far been limited to intrinsic material resonances[21] and the motion of trapped ions[8]. Entanglement of engineered (opto-)mechanical resonances, on the other hand, would provide a route towards scalable quantum networks. The freedom of designing and choosing optical resonances would allow operation in the entire frequency range of the technologically important C-, S- and L-bands of fibre-optic telecommunications. Together with dense wavelength-division multiplexing (on the ITU-T grid; ITU-T, International Telecommunication Union Standardization Sector), this could enable quantum nodes separated by long distances (about 100 km) that can communicate at large bandwidths. State-of-the-art engineered mechanical elements have energy lifetimes that typically range between micro-[15] and milliseconds[22], which would allow entanglement distribution on a regional level[23]. In addition, these entangled mechanical systems could be interfaced with microwaves[24], opening up the possibility of integrating superconducting quantum processors in the local nodes of the network.

Here we report on the observation of distributed entanglement between two nanomechanical resonators, mediated by telecommunication-wavelength photons. We use the DLCZ protocol[25], which was experimentally pioneered with ensembles of cold atoms[4]. The entanglement is generated probabilistically through the conditional preparation of a single phonon, heralded by the detection of a signal photon that could

originate from either of two identical optomechanical oscillators. Fabrication imperfections have previously limited the use of artificial structures, requiring external tuning mechanisms to render such systems indistinguishable. Here we demonstrate not only that obtaining sufficiently identical devices is in fact possible through nanofabrication, but also that our method could in principle be applied to more than two systems.

The mechanical oscillators that we use in our experiment are nano-structured silicon beams with co-localized mechanical and optical resonances. Radiation pressure forces and the photoelastic effect couple the optical and mechanical modes with a rate g_0, causing the optical frequency to shift under the displacement of the mechanical oscillator[26]. This effect can be used to selectively address Stokes and anti-Stokes transitions by driving the optical resonance with detuned laser beams, resulting in a linear optomechanical interaction. As was recently shown, this technique can be used to create non-classical mechanical and optomechanical states at the single-quantum level for individual devices by using photon counting and post-selection[15,19].

To apply the DLCZ scheme to the entanglement of two separate optomechanical crystals, a critical requirement is that the photons emitted from the optomechanical cavities must be indistinguishable. This can be achieved by creating a pair of nanobeams with identical optical and mechanical resonances. Until now, however, fabrication variations have inhibited the deterministic generation of identical devices and the design of current oscillators does not include any tuning capabilities. Considering the optical mode alone, typical fabrication runs result in a spread of the resonance frequency of about 2 nm around the centre wavelength. Therefore, finding a pair of matching optical resonances on two chips close to a target frequency currently relies on fabricating a large enough set, in which the probability of obtaining an identical pair is sufficiently high. In fact, this is achievable with a few hundred devices per chip (see Supplementary Information for details). In addition, a small mismatch in the mechanical frequencies, which is typically around 1%, can readily be compensated by appropriate manipulation of the optical pulse frequencies in the experiment.

For the experiments presented here, we chose a pair of devices with optical resonances at wavelength $\lambda = 1{,}553.8$ nm (optical quality factor $Q = 2.2 \times 10^5$ and $g_0/(2\pi) = 550$ kHz and 790 kHz for devices A and B, respectively; see Fig. 1). For these structures, the mechanical resonance frequencies are centred around $\Omega_m/(2\pi) \approx 5.1$ GHz and have a difference of $\Delta\Omega_m/(2\pi) = 45$ MHz. The two chips are mounted 20 cm apart in a dilution refrigerator. Although we use a single cryostat, there is in principle no fundamental or technical reason for keeping the devices in a common cold environment. For our setup, if the telecommunication fibres linking the two devices were to be unwrapped, our setup would already allow us to bridge a separation of about 70 m between the two chips without further modification.

The protocol[25] for the creation and verification of the remote mechanical entanglement consists of three steps (for a schematic,
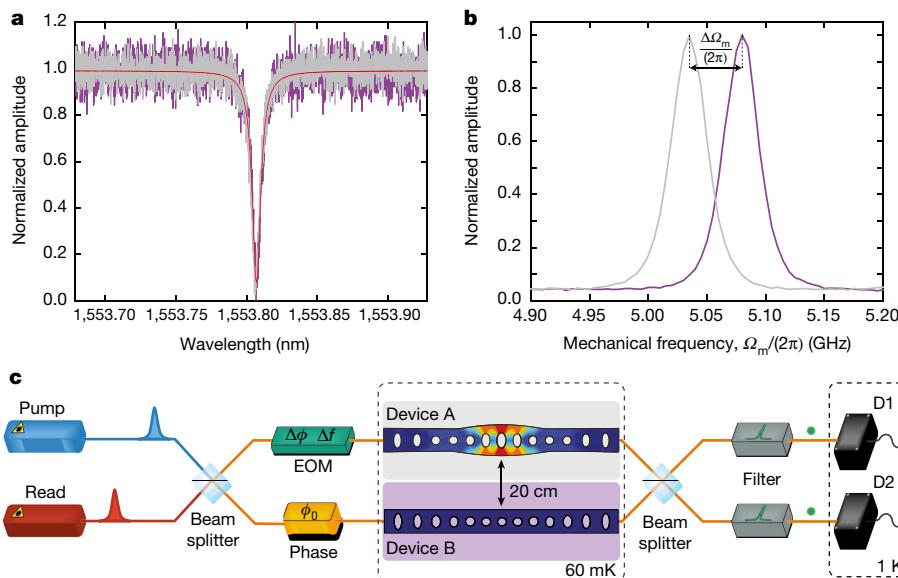
Fig. 1 | Devices and experimental setup. a, Optical resonances of device A (grey) and device B (magenta). The Lorentzian fit result (red line) yields a quality factor of $Q \approx 2.2 \times 10^5$ for each cavity. **b**, Mechanical resonances of device A (grey) and device B (magenta). The normalized mechanical resonances are measured through the optomechanical sideband scattering rates. The linewidth is limited by the bandwidth of the optical pulses and filters. The frequencies of the devices differ by $\Delta\Omega_m/(2\pi) = 45$ MHz, which could result in distinguishable photons, potentially reducing the entanglement in the system. We compensate for this shift by tuning the optical pump fields accordingly through serrodyning, erasing any information that could lead to a separable state. **c**, Experimental setup. We create optical pulses using two lasers, which are detuned to the Stokes (pump) and anti-Stokes (read) transition of the optomechanical cavities.

The lasers are then combined on a 50/50 beam splitter, which forms an interferometer with a second combining beam splitter. Each arm of the interferometer contains one of the mechanical oscillators, cooled to its ground state using a dilution refrigerator (central dashed rectangle). The phase of the interferometer, ϕ_0, is stabilized using a fibre stretcher (labelled 'phase'), while the phase difference between the pulses, $\Delta\phi$, is controlled using an electro-optic modulator (EOM). The same EOM is also used for serrodyning. Optical filters in front of two superconducting single-photon detectors (D1, D2) ensure that only photons scattered onto the cavity resonance are detected, whereas the original laser pulses are completely suppressed. The mechanical devices are physically separated by 20 cm and their optical separation is around 70 m.

see Fig. 2). First, the two mechanical resonators are cryogenically cooled, and thus initialized close to their quantum ground states[15,19,22] (see Supplementary Information). Second, a weak 'pump' pulse tuned to the upper mechanical sideband (at frequency $\omega_{pump} = 2\pi c/\lambda + \Omega_m$, where c is the speed of light), is sent into a phase-stabilized interferometer (with a fixed phase difference ϕ_0, see Fig. 1 and Supplementary Information) with one device in each arm. This drives the Stokes process—that is, the scattering of a pump photon into the cavity resonance while simultaneously creating a phonon[15]. The presence of a single phonon is heralded by the detection of a scattered Stokes photon in one of our superconducting nanowire single-photon detectors. The two optical paths of the interferometer are overlapped on a beam splitter, and a variable optical attenuator is set on one of the arms so that a scattered photon from either device is equally likely to reach either detector. The heralding detection event therefore contains no information about which device the scattering took place in and thus where the phonon was created. The energy of the pulse is tuned to ensure that the scattering probability $p_{pump} \approx 0.7\%$ is low, making the likelihood of simultaneously creating phonons in both devices negligible. The heralding measurement therefore projects the mechanical state into a superposition of a single-excitation state in device A ($|A\rangle = |1\rangle_A|0\rangle_B$) or device B ($|B\rangle = |0\rangle_A|1\rangle_B$), with the other device remaining in the ground state. The joint state of the two mechanical systems

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|1\rangle_A|0\rangle_B \pm e^{i\theta_m(0)}|0\rangle_A|1\rangle_B) \qquad (1)$$

is therefore entangled, where $\theta_m(0) = \phi_0$ is the phase with which the mechanical state is initialized at delay $\tau = 0$. This phase is determined from the relative phase difference that the pump beam acquires

in the two interferometer arms[4], which we can choose using our interferometer lock. However, because the two mechanical frequencies differ by $\Delta\Omega_m$, the phase of the entangled state will continue to evolve as $\theta_m(\tau) = \phi_0 + \Delta\Omega_m\tau$. The sign in equation (1) reflects which detector is used for heralding, with $+$ ($-$) corresponding to the positive (negative) detector, as defined by the sign convention of the interferometer phase ϕ_0.

In the third step of our protocol, we experimentally verify the entanglement between the two mechanical oscillators. To achieve this, we map the mechanical state onto an optical field using a 'read' pulse after a variable delay τ. This relatively strong pulse is tuned to the lower mechanical sideband of the optical resonance ($\omega_{read} = 2\pi c/\lambda - \Omega_m$). At this detuning, the field drives the anti-Stokes transition—that is, a pump photon is scattered onto the cavity resonance while annihilating a phonon[15]. Ideally, this state transfer will convert $|\Psi\rangle$ into

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|1\rangle_{r_A}|0\rangle_{r_B} \pm e^{i(\theta_r + \theta_m(\tau))}|0\rangle_{r_A}|1\rangle_{r_B}) \qquad (2)$$

where r_A and r_B are the optical modes in the two interferometer arms. The state of the optical field now contains the mechanical phase as well as the phase difference θ_r acquired by the read pulse. We can add an additional phase offset $\Delta\phi$ to the read pulse in one of the interferometer arms so that $\theta_r = \phi_0 + \Delta\phi$ by using an electro-optic phase modulator, as shown in Fig. 1. Sweeping $\Delta\phi$ allows us to probe the relative phase $\theta_m(\tau)$ between the superpositions $|A\rangle$ and $|B\rangle$ of the mechanical state for fixed delays τ. To avoid substantial absorption heating creating thermal excitations in the oscillators, we limit the energy of the read pulse to a state-swap fidelity of about 3.4%, reducing the number of added incoherent phonons to about 0.07 at a delay of $\tau = 123$ ns (see Supplementary Information).
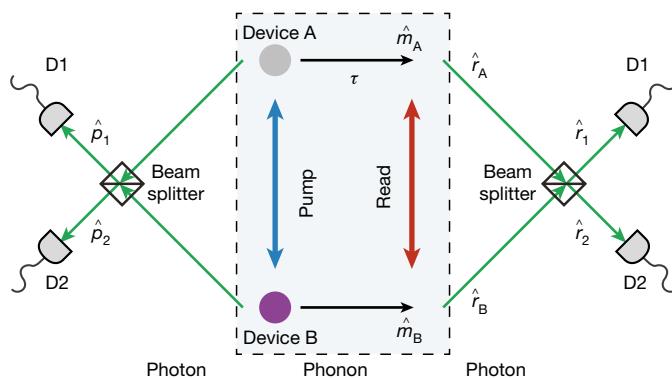
Fig. 2 | Creation and detection of entanglement between two remote mechanical oscillators. A pump pulse detuned to the Stokes sideband of two identical optomechanical resonators is sent into an interferometer, creating a single excitation in either device A or B. This process emits a photon on resonance with one of the cavities, and the two possible paths are superimposed using a beam splitter (black square) when exiting the interferometer (left). Detection of this photon in one of the single-photon detectors projects the two mechanical systems into an entangled state, in which neither device can be described separately. To verify this non-separable state, an optical read pulse tuned to the anti-Stokes sideband is sent into the interferometer with a delay of τ, de-exciting the mechanical systems and emitting another on-resonance photon into modes r_i ($i = A, B$) with operators $\hat{r}_i$. The two optical paths are again superimposed on the same beam splitter (right), and the photon is detected, allowing us to measure various second-order correlation functions, which are used to test an entanglement witness. The operators $\hat{p}_j$ and $\hat{r}_j$, with $j = 1, 2$, denote the optical modes created from the pump and the read pulses, respectively, after recombination on the beam splitter and $\hat{m}_i$ ($i = A, B$) are the operators of the mechanical modes. We note that in our experiment, the detectors used for the pump and read photons are identical (see Fig. 1).

So far we have neglected the consequence of slightly differing mechanical resonance frequencies for our heralding scheme. To compensate for the resulting frequency offset in the scattered (anti-) Stokes photons and to erase any available 'which device' information, we shift the frequency of the laser pulses by means of serrodyning (see Supplementary Information). Specifically, we use the electro-optic phase modulator, which controls the phase offset $\Delta\phi$, to also shift the frequency of the pump (read) pulses to device A by $+\Delta\Omega_m$ ($-\Delta\Omega_m$). The frequency differences of the pulses in the two opposing paths cancel out their mechanical frequency differences exactly, ensuring that the scattered photons at the output of the interferometer are indistinguishable.

To confirm that the measured state is indeed entangled, we need to distinguish it from all possible separable states, that is, the set of all states for which systems A and B can be described independently. A specifically tailored measure that can be used to verify this non-separability of the state is called an 'entanglement witness'. Here we use a witness that is designed for optomechanical systems[27]. In contrast to other path-entanglement witnesses based on partial state tomography, such as concurrence, this approach replaces measurements of third-order coherences, $g^{(3)}$, by expressing them as second-order coherences, $g^{(2)}$, assuming linear interactions between Gaussian states. This greatly simplifies the requirements and reduces the measurement times for our experiments. Because the coherences refer to the unconditional states, the nonlinear detection and state projection do not contradict these assumptions. The above assumptions are satisfied for our system because the initial mechanical states of our devices are in fact thermal states close to the corresponding quantum ground states (step 1 of our protocol; see Supplementary Information) and we use linearized optomechanical interactions (described in steps 2 and 3)[28]. The upper bound for this witness of mechanical entanglement is given by[27] (see Supplementary Information).

$$R_m(\theta, j) = 4 \frac{g^{(2)}_{r_1, p_j}(\theta) + g^{(2)}_{r_2, p_j}(\theta) - 1}{(g^{(2)}_{r_1, p_j}(\theta) - g^{(2)}_{r_2, p_j}(\theta))^2} \qquad (3)$$

in a symmetric setup. In equation (3), $\theta = \theta_r + \theta_m$, $j = 1, 2$ denotes the heralding detectors and $g^{(2)}_{r_i, p_j} = \langle \hat{r}_i^\dagger \hat{p}_j^\dagger \hat{r}_i \hat{p}_j \rangle / (\langle \hat{r}_i^\dagger \hat{r}_i \rangle \langle \hat{p}_j^\dagger \hat{p}_j \rangle)$ is the second-order coherence between the photons scattered by the pump pulse (with $\hat{p}_j^\dagger$ and $\hat{p}_j$ the creation and annihilation operators, respectively, of the mode going to detector j) and the converted phonons from the read pulse (with $\hat{r}_j^\dagger$ and $\hat{r}_j$ the creation and annihilation operators, respectively, of the mode going to detector j). For all separable states of the mechanical oscillators A and B, the witness yields $R_m(\theta, j) \geq 1$ for any θ and j. Hence, if there exists a θ and j for which $R_m(\theta, j) < 1$, the mechanical systems must be entangled.

Although entanglement witnesses are designed to be efficient classifiers, they typically depend on the individual characteristics of the experimental setup. If, for example, the second beam splitter (see Fig. 1) were to malfunction and act as a perfect mirror—that is, if all photons from device A (B) were transmitted to detector 1 (2)—then $R_m(\theta, j)$ could still be less than 1 for separable states. This is because the witness in equation (3) estimates the visibility of the interference between $|A\rangle$ and $|B\rangle$ from a single measurement, without requiring a full phase scan of the interference fringe. To ensure the applicability of the witness, we therefore verify experimentally that our system fulfils its assumptions. We first check whether our setup is balanced by adjusting the energy of the pump pulses in each arm, as described above. This guarantees that the scattered photon fluxes impinging on the beam splitter from both arms are equal (see Supplementary Information). To make the detection symmetric, we use heralding detection events from both superconducting nanowire single-photon detectors—that is, we obtain the actual bound on the entanglement witness $R_{m,sym}(\theta)$ from averaging measurements of $R_m(\theta, 1)$ and $R_m(\theta, 2)$ (see Supplementary Information). By choosing a phase θ such that the correlations between different detectors exceed the correlations at the same detector, $g^{(2)}_{r_i, p_j, i \neq j} > g^{(2)}_{r_i, p_i}$ with $i, j \in \{1, 2\}$, we avoid our measurements' susceptibility to unequal splitting ratios applied by the beam splitter.

In Fig. 3, we show a series of measurements of the second-order coherence $g^{(2)}$, performed by sweeping $\Delta\phi$ with a readout delay of $\tau = 123$ ns, which verify the coherence between $|A\rangle$ and $|B\rangle$. Using these data, we chose an optimal phase setting $\theta = \theta_{opt}$ with $\Delta\phi = 0.2\pi$ for the main experiment. We obtain $R_{m,sym}(\theta_{opt}) = 0.74^{+0.12}_{-0.06}$, which is well below the separability bound of 1. By including measurements at the non-optimal adjacent phases $\Delta\phi = 0$ and 0.25π, the statistical uncertainty improves, and we obtain $R_{m,sym}([\theta_{opt} - 0.2\pi, \theta_{opt} + 0.05\pi]) = 0.74^{+0.08}_{-0.05}$. Hence, we experimentally observe entanglement between the two remote mechanical oscillators with a confidence level above 99.8%.

The coherence properties of the generated state can be characterized through the decay of the visibility

$$V = \frac{\max(g^{(2)}_{r_i, p_j}) - \min(g^{(2)}_{r_i, p_j})}{\max(g^{(2)}_{r_i, p_j}) + \min(g^{(2)}_{r_i, p_j})} \qquad (4)$$

We therefore sweep the delay time τ between the pump pulse and the read pulse. The mechanical frequency difference $\Delta\Omega_m$ allows us to sweep a full interference fringe by changing the delay τ by 22 ns. Owing to the technically limited hold time of our cryostat, this sweep had to be performed at a higher bath temperature of about $80-90$ mK (see Fig. 1), yielding a slightly lower, thermally limited visibility at short delays when compared to the data in Fig. 3. By varying the delay further, we observe interference between $|A\rangle$ and $|B\rangle$ ($V > 0$) up to $\tau \approx 3$ μs (see Fig. 4). The loss of coherence can be explained by absorption heating and mechanical decay (see Supplementary Information) and appears to be limited at long delays τ by the lifetime $1/\Gamma_A \approx 4$ μs of device A, which has the shorter lifetime of the two devices.
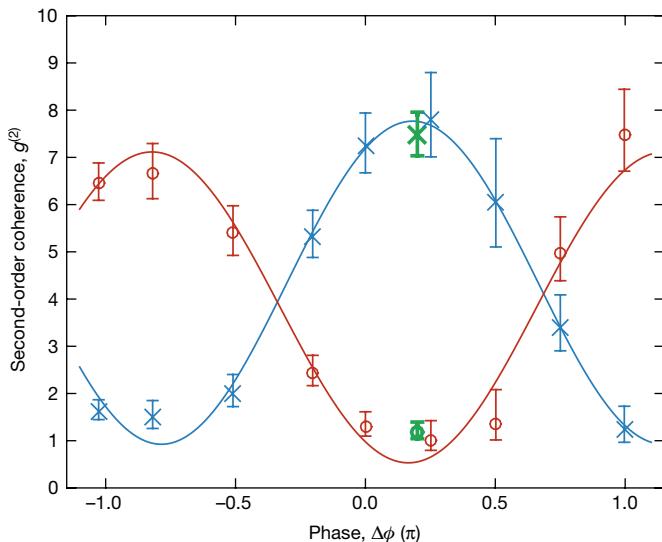
Fig. 3 | Phase sweep of the entangled state. We vary the phase difference between the pump and the read pulses, $\Delta\phi$, and measure the second-order coherence $g^{(2)}$ of the Raman-scattered photons for a fixed delay of $\tau = 123$ ns between the pulses. Blue crosses represent measurements of $g^{(2)}_{r_i,p_j,i\neq j}$ and red circles are the results for $g^{(2)}_{r_i,p_i}$, where $i,j \in \{1,2\}$. We fit simple sine functions (shown as solid lines) to each of the datasets as guides to the eye. The sinusoidal dependence on the phase clearly highlights the coherence of the entangled mechanical state. We observe a periodicity of 1.95π, in good agreement with the expected value of 2π for single-particle interference (see equation (2))[27]. The phase sweep allows us to identify the optimal phase $\Delta\phi = 0.2\pi$ for maximum visibility, at which we acquire additional data (green cross and circle) to determine the entanglement witness with sufficient statistical significance. All error bars represent a 68% confidence interval.

Fig. 4 | Time sweep of the entangled state. Shown is the interference of the entangled mechanical state at different delays τ between the pump and read pulses, with the phase of the interferometer, ϕ_0, and the phase difference between the pump and read pulses, $\Delta\phi$, fixed. The blue crosses represent the measurements of $g^{(2)}_{r_i,p_j,i\neq j}$ and red circles are the results for $g^{(2)}_{r_i,p_i}$, where $i,j \in \{1,2\}$. The solid lines are sinusoidal fits averaged over the two out-of-phase components for each delay window and serve as a guide to the eye. The coherence of the entangled state is reduced over time, which can be seen by the decay of the interference visibility (inset). This decoherence is consistent with a delayed optical absorption heating and the mechanical decay time of about 4 μs of device A. The inset shows the visibility of the interference (green crosses) and the expected upper bound on the visibility due to heating and mechanical decay (orange line; see Supplementary Information). All error bars represent a 68% confidence interval.

We have experimentally demonstrated entanglement between two engineered mechanical oscillators separated spatially by 20 cm and optically by 70 m. Imperfections in the fabrication process and the resulting small deviations of optical and mechanical frequencies for nominally identical devices are overcome through the statistical selection of devices and optical frequency shifting using a serrodyne approach. The mechanical systems do not interact directly at any point, but are interfaced remotely through optical photons in the telecommunication-wavelength band. The coherence time of the entangled state is several microseconds and appears to be limited by the mechanical lifetime of the devices and by absorption heating. Both of these limitations can be considerably mitigated. On the one hand, optical absorption can be substantially suppressed by using intrinsic, desiccated silicon[29]. Mechanical lifetimes, on the other hand, can be greatly increased by adding a phononic bandgap shield[22]. Although our devices are engineered to have short mechanical lifetimes[19,30], earlier designs including such a phononic shield have reached[22] $1/\Gamma \approx 0.5$ ms and could still be further improved. Combined with reduced optical absorption, which would allow efficient laser cooling, such lifetimes can potentially put our devices on par with other state-of-the-art quantum systems[31].

Our experiment demonstrates a protocol for realistic, fibre telecommunication-compatible entanglement distribution using engineered mechanical quantum systems. With the current parameters of our system, a device separation of 75 km using commercially available telecommunication fibres would result in a drop of less than 5% in the interference visibility (see discussion in Supplementary Information for more details). The system presented here is directly scalable to include more devices (see Supplementary Information) and could be integrated into a real quantum network. Combining our results with those of optomechanical devices capable of transferring quantum information from the optical to the microwave domain, which is a highly active field of research[24,32,33], could provide a backbone for a future quantum internet based on superconducting quantum computers.

Data availability

All relevant data generated and analysed during this study are included in this paper (and its Supplementary Information).

1. Kimble, H. J. The quantum internet. *Nature* **453**, 1023–1030 (2008).
2. Jensen, K. et al. Quantum memory for entangled continuous-variable states. *Nat. Phys.* **7**, 13–16 (2011).
3. Reim, K. F. et al. Single-photon-level quantum memory at room temperature. *Phys. Rev. Lett.* **107**, 053603 (2011).
4. Chou, C. W. et al. Measurement-induced entanglement for excitation stored in remote atomic ensembles. *Nature* **438**, 828–832 (2005).
5. Matsukevich, D. N. et al. Entanglement of remote atomic qubits. *Phys. Rev. Lett.* **96**, 030405 (2006).
6. Ritter, S. et al. An elementary quantum network of single atoms in optical cavities. *Nature* **484**, 195–200 (2012).
7. Moehring, D. L. et al. Entanglement of single-atom quantum bits at a distance. *Nature* **449**, 68–71 (2007).
8. Jost, J. D. et al. Entangled mechanical oscillators. *Nature* **459**, 683–685 (2009).
9. Usmani, I. et al. Heralded quantum entanglement between two crystals. *Nat. Photon.* **6**, 234–237 (2012).
10. Saglamyurek, E. et al. Quantum storage of entangled telecom-wavelength photons in an erbium-doped optical fibre. *Nat. Photon.* **9**, 83–87 (2015).
11. Hensen, B. et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
12. Teufel, J. D. et al. Sideband cooling of micromechanical motion to the quantum ground state. *Nature* **475**, 359–363 (2011).
13. Chan, J. et al. Laser cooling of a nanomechanical oscillator into its quantum ground state. *Nature* **478**, 89–92 (2011).
14. Palomaki, T. A., Teufel, J. D., Simmonds, R. W. & Lehnert, K. W. Entangling mechanical motion with microwave fields. *Science* **342**, 710–713 (2013).
15. Riedinger, R. et al. Non-classical correlations between single photons and phonons from a mechanical oscillator. *Nature* **530**, 313–316 (2016).
16. Wollman, E. E. et al. Quantum squeezing of motion in a mechanical resonator. *Science* **349**, 952–955 (2015).
17. O'Connell, A. D. et al. Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).

18. Chu, Y. et al. Quantum acoustics with superconducting qubits. *Science* **358**, 199–202 (2017).
19. Hong, S. et al. Hanbury Brown and Twiss interferometry of single phonons from an optomechanical resonator. *Science* **358**, 203–206 (2017).
20. Reed, A. P. et al. Faithful conversion of propagating quantum information to mechanical motion. *Nat. Phys.* **13**, 1163–1167 (2017).
21. Lee, K. C. et al. Entangling macroscopic diamonds at room temperature. *Science* **334**, 1253–1256 (2011).
22. Meenehan, S. M. et al. Pulsed excitation dynamics of an optomechanical crystal resonator near its quantum ground state of motion. *Phys. Rev. X* **5**, 041002 (2015).
23. Razavi, M., Piani, M. & Luotkenhaus, N. Quantum repeaters with imperfect memories: cost and scalability. *Phys. Rev. A* **80**, 032301 (2009).
24. Bochmann, J., Vainsencher, A., Awschalom, D. D. & Cleland, A. N. Nanomechanical coupling between microwave and optical photons. *Nat. Phys.* **9**, 712–716 (2013).
25. Duan, L. M., Lukin, M. D., Cirac, J. I. & Zoller, P. Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414**, 413–418 (2001).
26. Chan, J. *Laser Cooling of an Optomechanical Crystal Resonator to its Quantum Ground State of Motion*. Ph.D. thesis, California Institute of Technology (2012).
27. Børkje, K., Nunnenkamp, A. & Girvin, S. M. Proposal for entangling remote micromechanical oscillators via optical measurements. *Phys. Rev. Lett.* **107**, 123601 (2011).
28. Wieczorek, W. et al. Optimal state estimation for cavity optomechanical systems. *Phys. Rev. Lett.* **114**, 223601 (2015).
29. Asano, T., Ochi, Y., Takahashi, Y., Kat-suhiro, K. & Noda, S. Photonic crystal nanocavity with a Q factor exceeding eleven million. *Opt. Express* **25**, 1769–1777 (2017).
30. Patel, R. N., Sarabalis, C. J., Jiang, W., Hill, J. T. & Safavi-Naeini, A. H. Engineering phonon leakage in nanomechanical resonators. *Phys. Rev. Appl.* **8**, 041001 (2017).
31. Maring, N. et al. Photonic quantum state transfer between a cold atomic gas and a crystal. *Nature* **551**, 485–488 (2017).
32. Rueda, A. et al. Efficient microwave to optical photon conversion: an electro-optical realization. *Optica* **3**, 597–604 (2016).
33. Higginbotham, A. P. et al. Electro-optic correlations improve an efficient mechanical converter. Preprint at https://arxiv.org/abs/1712.06535 (2017).

LETTER

Stabilized entanglement of massive mechanical oscillators

C. F. Ockeloen-Korppi[1], E. Damskägg[1], J.-M. Pirkkalainen[1], M. Asjad[2], A. A. Clerk[3], F. Massel[2], M. J. Woolley[4] & M. A. Sillanpää[1]*

Quantum entanglement is a phenomenon whereby systems cannot be described independently of each other, even though they may be separated by an arbitrarily large distance[1]. Entanglement has a solid theoretical and experimental foundation and is the key resource behind many emerging quantum technologies, including quantum computation, cryptography and metrology. Entanglement has been demonstrated for microscopic-scale systems, such as those involving photons[2–5], ions[6] and electron spins[7], and more recently in microwave and electromechanical devices[8–10]. For macroscopic-scale objects[8–14], however, it is very vulnerable to environmental disturbances, and the creation and verification of entanglement of the centre-of-mass motion of macroscopic-scale objects remains an outstanding goal. Here we report such an experimental demonstration, with the moving bodies being two massive micromechanical oscillators, each composed of about 10^{12} atoms, coupled to a microwave-frequency electromagnetic cavity that is used to create and stabilize the entanglement of their centre-of-mass motion[15–17]. We infer the existence of entanglement in the steady state by combining measurements of correlated mechanical fluctuations with an analysis of the microwaves emitted from the cavity. Our work qualitatively extends the range of entangled physical systems and has implications for quantum information processing, precision measurements and tests of the limits of quantum mechanics.

There exist several proposals for how cavity optomechanical setups could be used to entangle the motional quantum states of two massive mechanical oscillators[15,18–22]. In such setups, two movable mirrors are incorporated into a resonant optical cavity, and radiation pressure forces inside the cavity can be tailored so that the motion of the mirrors becomes highly correlated and even entangled. This approach does not require any direct interaction between the moving masses, allowing them to be spatially separated. A correlated reduction of noise below the thermal level in the motion of two mechanical oscillators has previously been demonstrated experimentally[23,24], but not near the quantum level, as required for entanglement.

Our work is based on a series of proposals[15,20,21] for using reservoir engineering to stabilize two cavity-coupled mechanical oscillators into a steady state that is entangled. The methods that we use are extensions of an approach used in recent work to 'squeeze' the motion of a single oscillator[25–27]. An oscillator with frequency ω_1 and position operator $x_1(t) = X_1(t)\cos(\omega_1 t) + P_1(t)\sin(\omega_1 t)$ is squeezed if the variance of either of the quadrature-amplitude operators X_1 and P_1 is smaller than the quantum zero-point fluctuation level. Introducing a second oscillator with frequency ω_2 and position operator x_2 expressed in terms of quadratures as above, one can introduce four collective quadrature operators, $X_{\pm} = \frac{1}{\sqrt{2}}(X_2 \pm X_1)$ and $P_{\pm} = \frac{1}{\sqrt{2}}(P_2 \pm P_1)$. The state corresponding to the variances of either X_+ and P_-, or X_- and P_+, being reduced below the quantum zero-point fluctuation level is a canonical entangled state known as the two-mode squeezed state. Specifically, the state is entangled if $\langle X_+^2 \rangle + \langle P_-^2 \rangle < 1$ (see Fig. 1e), a criterion commonly referred to as the Duan inequality[28]. Although such Gaussian states exhibit a

positive definite Wigner function, they can in principle have arbitrarily large amounts of entanglement, can lead to violations of local realism[29,30], and are essentially the state considered in the Einstein–Podolsky–Rosen gedanken experiment[1].

In our experiment, a single driven cavity mode is used both to prepare a correlated state of two mechanical oscillators and to directly measure fluctuations in the X_+ collective quadrature via a two-mode back-action evading (BAE) measurement[15–17]; we find $\langle X_+^2 \rangle \approx 0.41 \pm 0.04$. Although a direct measurement of P_- is not possible, by constraining the system parameters and analysing the full output spectrum of our cavity, we can infer its fluctuations, $\langle P_-^2 \rangle \approx 0.42 \pm 0.08$ (see, for example, Wollman et al.[25]). Hence, we have an entangled state with $\langle X_+^2 \rangle + \langle P_-^2 \rangle \approx 0.83 \pm 0.13 < 1$ of the centre-of-mass motion of two oscillators, each consisting of approximately 10^{12} atoms.

As shown in Fig. 1a, we use a microwave-frequency realization of cavity optomechanics involving two micromechanical drum oscillators[31] and a superconducting on-chip circuit acting as the electromagnetic cavity (frequency ω_c). The oscillators' positions affect the total capacitance and thus modulate the frequency of the cavity. This creates an effective radiation pressure interaction similar to that between photons and mirrors in an optical cavity. To generate two-mode squeezing and entanglement, we pump the system with two strong pump microwave tones at frequencies ω_- and ω_+, below and above the cavity frequency, respectively, as shown in Fig. 1b, c.

To describe this system, we initially assume for simplicity that the two oscillators have equal single-photon radiation-pressure coupling strengths, g_0, and that the pump tones are applied at the red- and blue-sideband frequencies $\omega_- = \omega_c - \omega_1$ and $\omega_+ = \omega_c + \omega_2$, respectively. Details of the theoretical model, including non-idealities, are discussed in Supplementary Information. The strong pump tones enhance the radiation-pressure interaction, yielding many-photon coupling rates $G_{\pm} = g_0 \alpha_{\pm}$, where $\alpha_{\pm}$ are the field amplitudes induced in the resonator by the pump tones at $\omega_{\pm}$. We also introduce the mechanical Bogoliubov modes, which are obtained by a two-mode squeezing transformation on the original mechanical annihilation operators, that is, $\beta_1 = b_1 \cosh r + b_2^\dagger \sinh r$ and $\beta_2 = b_2 \cosh r + b_1^\dagger \sinh r$, where r is the two-mode squeezing parameter, given by $\tanh r = G_+/G_-$. By defining $\Omega = (\omega_2 - \omega_1)/2$ and working in a rotating frame (at frequency $\omega_c + \Omega$ for the cavity and $(\omega_2 + \omega_1)/2$ for each mechanical oscillator), the linearized optomechanical Hamiltonian becomes:

$$H = -\Omega a^\dagger a + \Omega(\beta_2^\dagger \beta_2 - \beta_1^\dagger \beta_1) + \mathcal{G}[a^\dagger(\beta_1 + \beta_2) + a(\beta_1^\dagger + \beta_2^\dagger)] \quad (1)$$

where $\mathcal{G} = \sqrt{G_-^2 - G_+^2}$ is the effective optomechanical coupling rate. This Hamiltonian is essentially that used in the study of Woolley and Clerk[15], but with the pump frequencies set as in the system of Wang and Clerk[21]. It describes the cavity cooling of the Bogoliubov modes towards their mutual ground state, which corresponds to a stabilized, two-mode squeezed state of the bipartite mechanical system. In contrast to

---

[1]Department of Applied Physics, Aalto University, Aalto, Finland. [2]Department of Physics and Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland. [3]Institute for Molecular Engineering, University of Chicago, Chicago, IL, USA. [4]School of Engineering and Information Technology, UNSW Canberra, Canberra, Australian Capital Territory, Australia. *e-mail: mika.sillanpaa@aalto.fi

dynamical protocols[10,13] the system stays entangled indefinitely. Here, the use of non-degenerate mechanical frequencies is essential to ensure that both Bogoliubov modes are efficiently cooled by different frequency components of the cavity. Asymmetries in the single-photon couplings g_1 and g_2 of the two oscillators, or non-zero detunings of the pump tones (by amounts $\delta_\pm$, see Fig. 1b) from the mechanical sidebands, introduce additional terms in the Hamiltonian of equation (1), and one might expect the steady-state entanglement to be reduced. However, we find numerically that we can greatly compensate for asymmetries in couplings by optimizing the pump detunings $\delta_\pm \neq 0$.

As shown in Fig. 1d, an essential part of the entanglement verification strategy consists of two-mode BAE detection[16,17] operated in the same cavity mode, allowing mapping of the mechanical motion to the output field. This involves two relatively weak probe tones applied at $\omega_{d\pm} \approx \omega_c \pm (\omega_2 + \omega_1)/2$, approximately in the middle of the sideband frequencies. To preserve the same rotating frame for the creation and detection of the two-mode squeezing, we strictly require $\omega_{d+} - \omega_{d-} = \omega_+ - \omega_-$, ideally up to complete phase coherence between the tones. Similarly to the pump tones, the probe tones induce effective couplings $g_\pm = g_0 \alpha_\pm^d$ with amplitudes $\alpha_\pm^d$, which are equal in the ideal two-mode BAE case. Because we are using the same cavity mode for both creating the entanglement via the pump tones and detecting it, the pump spectra and the probe spectra need to be independent. This is achieved by ensuring that the mechanical contributions to the output cavity spectrum from the pump and probe tones have negligible spectral overlap. Therefore, the faithful reconstruction of the X_+ collective quadrature spectrum from the probe signal is possible. In contrast to BAE detection of single-mode squeezing[25], both the pump and probe tones can be set to have optimal frequencies for the creation and detection of two-mode squeezing (see Fig. 1b–d).

In our device the two oscillators are separated by 600 μm, they have no direct coupling and the system is well described by equation (1). We use the fundamental drum modes of the oscillators with resonance frequencies $\omega_1/(2\pi) \approx 10.0$ MHz and $\omega_2/(2\pi) \approx 11.3$ MHz and linewidths $\gamma_1/(2\pi) \approx 106$ Hz and $\gamma_2/(2\pi) \approx 144$ Hz, respectively. The microwave cavity, with a frequency of $\omega_c/(2\pi) \approx 5.5$ GHz, has separate input and output ports. All the input signals are applied through a port coupled weakly at a rate of $\kappa_{Ei}/(2\pi) \approx 60$ kHz, whereas the output is strongly coupled at $\kappa_{Eo}/(2\pi) \approx 1.13$ MHz. The cavity has also internal losses with a rate of $\kappa_i/(2\pi) \approx 190$ kHz, and the sum of all the loss channels give a total linewidth of $\kappa/(2\pi) \approx 1.38$ MHz. We find that our fabrication process can produce basically identical single-photon couplings, $g_1/g_2 \approx 0.98$, for two oscillators of different frequencies; in fact, this is more than sufficient for generating entanglement because we find numerically that an asymmetry of up to about 20% can be compensated using detuning.

The motion of the mechanical oscillators is determined using the power scattered from the applied microwave (pump and probe) tones by their interaction with the oscillators. We collect this weak signal using standard techniques, including the use of a low-noise cryogenic microwave amplifier, followed by room-temperature signal analysis. Performing a sequence of calibrations, described in detail in Supplementary Information, is important for the experiment. First, using a standard thermal calibration with a single red-detuned tone, the mechanical modes are found to thermalize to equilibrium phonon occupation numbers of $n_1^T \approx 41$ and $n_2^T \approx 30$ for oscillators 1 and 2, respectively, at the base temperature of the dry dilution refrigerator, about 14 mK. These values give the initial variances of the collective quadratures $\langle X_\pm^2 \rangle^T$ and $\langle P_\pm^2 \rangle^T \approx 36$, where the superscript refers to the temperature T.

Second, we apply standard sideband cooling to each mechanical oscillator separately using a single red-detuned pump (see Fig. 2 for data from oscillator 1). This allows the characterization of the behaviour of the system under intense pumping. Importantly, it enables us to calibrate the gain of the detection system for the later interpretation of the spectrum under two-tone pumping, as well as the effective coupling of the red-detuned tone. The goal of calibrating the probe tones is to use
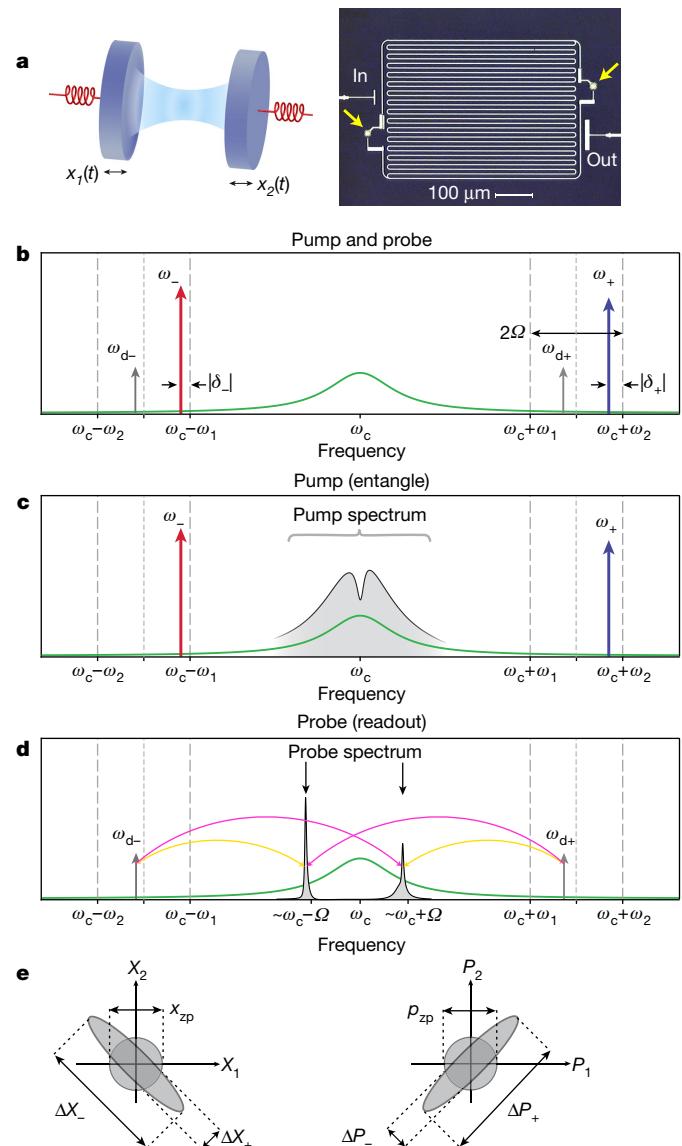
Fig. 1 | Creating and detecting motional entanglement. a, Schematic of two vibrating mirrors (frequencies ω_1, ω_2) coupled via an electromagnetic cavity with frequency ω_c (left), and a micrograph showing the microwave optomechanical device consisting of a superconducting transmission-line resonator, whose opposite ends are connected by two mechanical drum-type oscillators marked with arrows (right). The device is fabricated from aluminium, on a quartz chip. **b**, Spectrum of the microwave frequencies involved, showing the pump tones at frequencies $\omega_+ = \omega_c + \omega_2 + \delta_+$ and $\omega_- = \omega_c - \omega_1 + \delta_-$ (blue and red, respectively) and the probe tones with frequencies $\omega_{d\pm}$ (grey). Here, $\Omega = (\omega_2 - \omega_1)/2$. The bare-cavity response function is illustrated in green. **c**, The strong pump tones applied at frequencies $\omega_\pm$ create all-mechanical entanglement and carry information as incoherently scattered microwave light (pump spectrum). **d**, Two additional weak probe tones are applied at frequencies $\omega_{d\pm}$ to reconstruct the collective mechanical quadrature X_+ using a two-mode BAE measurement (probe spectrum). The curved arrows indicate the sideband processes that scatter phonons of frequencies ω_1 (yellow) or ω_2 (pink). **e**, Correlations in two-mode squeezing, shown in terms of fluctuations (shaded) of the quadrature amplitudes. Left, the sum of the X quadratures of the two oscillators fluctuates less than the zero-point level x_{zp} (circle). Right, the difference between P quadratures is similarly localized below the zero-point fluctuation level p_{zp}.

the total power in the probe spectra as a straightforward thermometer for the quadratures. Similarly to the case of the single red tone, we run a thermal calibration with both probe tones on, which allows us to determine the collective occupation number at a small probe power.
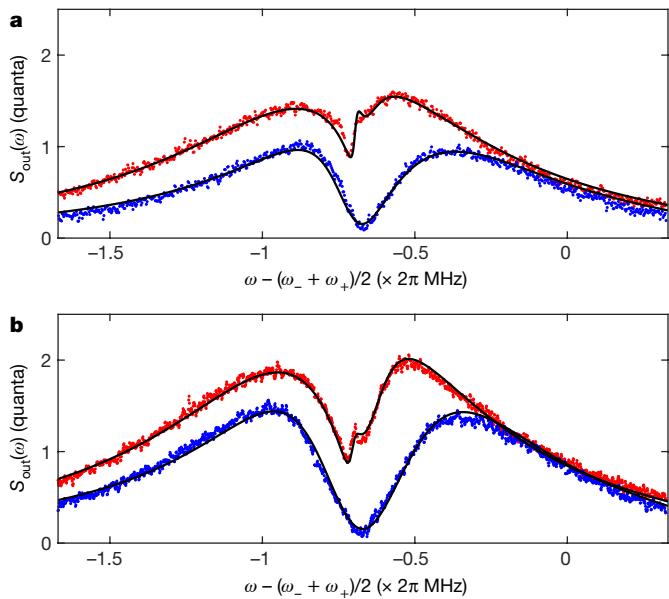
Fig. 2 | Pump spectra under two-tone driving. a, The pump amplitudes are $G_-/(2\pi) \approx 278\,$kHz and $G_+/(2\pi) \approx 166\,$kHz (dataset A, red points). **b**, Dataset B (red points), obtained with higher pump powers, $G_-/(2\pi) \approx 332\,$kHz and $G_+/(2\pi) \approx 210\,$kHz. The blue points, shown for reference, are the sideband-cooling calibration run results for oscillator 1. Theoretical curves are given by the solid lines. The remaining parameters are listed in Supplementary Information.

Subsequently, a power-sweep calibration of the probe tones associates a given signal strength with the quadrature variance.

Next we discuss the main experiment, which uses two pairs of tones, namely, the pump and probe tones. The pump tones are used to create entanglement, and the probe tones enable the measurement of the quadrature variances of the collective mechanical state. First, we focus on the spectrum of the pump tones. In Fig. 2, we display the pump output spectra $S_{out}(\omega)$ obtained from the cavity under several pumping conditions, expressed in absolute units (quanta) determined from the gain calibration. This spectrum contains information about the entanglement of the two oscillators because it allows us to infer the effective temperatures of the reservoirs associated with the cavity and the two oscillators. Our theoretical modelling uses standard input–output theory for electromagnetic cavities, treating the pump and probe tones as belonging to independent modes. For a given set of parameters, the variances of the collective quadratures can be evaluated within the same framework.

We now consider the spectrum resulting from the probe tones. The ability to control the relative phase ϕ of the two probe tones allows us to infer the variance of a general collective quadrature $X_+^\phi = X_+ \cos\phi + P_+ \sin\phi$. When written in terms of the spectra $S_{X_+}(\omega)$ and $S_{P_+}(\omega)$ of X_+ and P_+, respectively, the measured spectrum is then proportional to $S_{X_+^\phi}(\omega) = S_{X_+}(\omega)\cos^2\phi + S_{P_+}(\omega)\sin^2\phi$. In the measurement, the probe signal is visible as peaks on top of the pump spectrum at frequencies of about $\omega_c \pm \Omega$ on either side of the cavity frequency. To obtain the probe spectrum (Fig. 1d), we subtract the background measured in the absence of probe tones. In Fig. 3 we display the measured probe spectra $S_{out}^d(\omega)$ corresponding to dataset A. The probe power at $g_\pm/(2\pi) \approx 40\,$kHz was kept much smaller than the pump power. The theoretical model, calculated with the same parameters for all curves, is in excellent agreement with the experiment, including the positions and markedly non-Lorentzian lineshapes of the peaks, all of which depend on the phase. These unusual shapes arise because the two oscillators are pumped in an unequal manner (resonantly either on the red- or blue-sideband frequency), and they exhibit individual optical springs that add up to the collective spectrum. We infer the quadrature variance $\langle (X_+^\phi)^2 \rangle$ from the total integrated area of
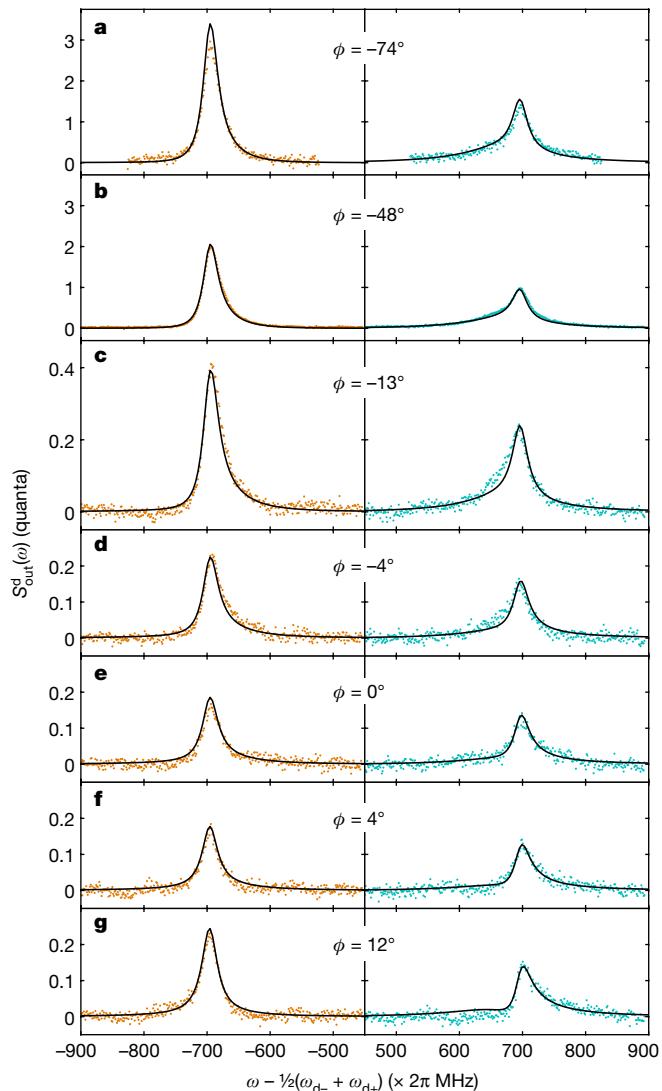
Fig. 3 | Two-mode BAE readout. Probe output spectra, recorded with the same parameters as the pump spectra in Fig. 2a (dataset A), that is, $G_-/(2\pi) \approx 278\,$kHz and $G_+/(2\pi) \approx 166\,$kHz. The left and right peaks correspond to the respective peaks in Fig. 1d. The probe phases are written in the panels, and the solid lines are theoretical predictions. **a, b**, Hot quadratures. **c**, Intermediate. **d–g**, Two-mode squeezing below the quantum zero-point fluctuation level.

the peaks. In contrast to approaches relying on the shapes of the peaks, this method is insensitive to the phase drift of the microwave sources that occurs during data acquisition. Within a typical integration time for one curve, approximately 30 min, ϕ can drift by several degrees, leading to slight departures from the theoretically determined curves.

From our modelling (see Supplementary Information), we confirm that the probe peak area faithfully reproduces the quadrature variances following our calibration. This strictly holds if the probe powers are perfectly matched, that is, $\alpha_+^d = \alpha_-^d$. This condition was calibrated without the pump tones being applied. However, for the theoretical fits in Fig. 3, we need to assume an imbalance of $g_-/g_+ \approx 1.055$, which we attribute to a shift of the cavity frequency. This shift occurs when the strong pump tones are present, causing a reduction in the blue-sideband probe tone in a system involving an input filter with a steep slope on the blue side. Given that g_- stays constant, the reduced g_+ means that using the probe area underestimates the quadrature variance, here by about 18%, as obtained numerically from the model. Therefore, we scale up the quadrature variances inferred from the probe areas by this percentage. A low blue-sideband pump power, similar to the best estimate, is observed in the pump spectra as well, but the
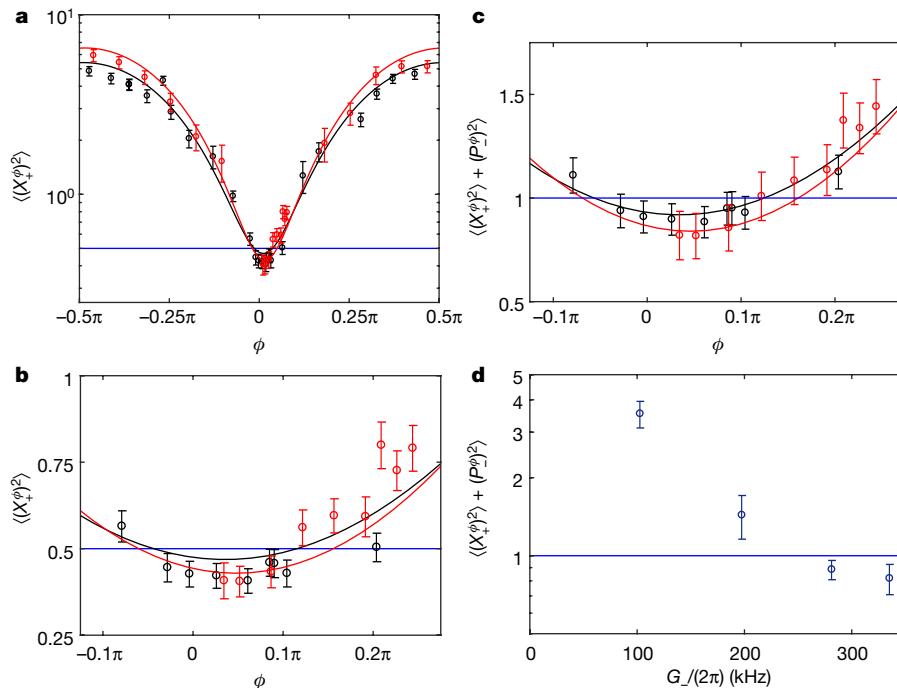
Fig. 4 | Fluctuations of collective quadratures. a, The X_+^ϕ quadrature variance, determined from the probe spectra. The black circles represent dataset A and the red circles correspond to dataset B. **b**, Magnification of **a**. **c**, The Duan quantity for entanglement as a function of the probe tone phase ϕ. **d**, The Duan quantity for the optimal value of ϕ, as a function of the strength of the red-detuned pump tone. The black and red solid lines are theoretical fits to the corresponding datasets, obtained using the bath temperatures determined by the pump spectra. The blue horizontal line marks the quantum zero-point fluctuations level. The error bars denote statistical confidence of two standard deviations.

red-sideband pump matches the calibration results, supporting our approach to handing this imbalance. We emphasize that the probe power calculation does not assume anything about the mechanical oscillators or the dynamics induced by the pump tones, but only assumes an understanding of the dynamics associated with the probe tones.

In Fig. 4a and b we display a measurement of the X_+^ϕ quadrature variance, with 95% statistical confidence intervals. In the optimal case of $\phi \approx 4°$, by minimizing the variance in Fig. 3, we obtain $\langle (X_+^\phi)^2 \rangle \approx 0.41 \pm 0.04$; that is, 0.9 dB below the vacuum noise. Several points fall well below the quantum zero-point noise level in both datasets A and B. Because the best theoretical fit to the measured probe spectra is obtained with dataset A, we base our main claims on these data. In dataset B, the theoretical fit shows good agreement with the measured pump spectrum although we believe that the probe spectrum was subject to larger phase drifts during the data acquisition.

Now we consider the measurement of the variance of the P_- quadrature, which is needed for examining the Duan criterion and verifying quantum entanglement. As mentioned, two-mode BAE probe detection does not couple to P_- or X_-. We therefore use the other source of information available—the pump spectrum—and combine it with the information provided by the probe detection. The variances are evaluated using a least-squares fit to an analytical expression describing the pump spectrum, using the three bath temperatures as adjustable parameters, combined with the aforementioned calibrations. The fits are shown in Fig. 2, displaying excellent agreement with the experiment. For dataset A, we obtain the variance $\langle P_-^2 \rangle \approx 0.45 \pm 0.08$. For the X_+ quadrature, we similarly get $\langle X_+^2 \rangle \approx 0.46 \pm 0.08$, close to the value obtained from the direct BAE detection method described above. Given our knowledge of the system parameters and the dynamics of this scheme, the two quadratures are expected to have variances within 5% of one another (see Supplementary Information), providing additional, BAE-based evidence for the value of $\langle P_-^2 \rangle$.

The error analysis of the probe measurement uses straightforward error propagation of the experimental calibration errors and of the statistical error from integrating the probe peak area. For the pump

spectrum, the analysis is complicated because it involves more parameters, some of which can considerably affect the steady-state entanglement. Here we adopt an error analysis method known as the Bayesian Monte Carlo method, similar to that used in Wollman et al.[25], to infer the parameters of the system, including uncertainties and correlations. This method generates a sample of the parameter distribution for which the theoretical model agrees with the measured pump spectra within the statistical uncertainty. We sample the posterior distributions of all parameters, and use them to estimate the confidence limits of the P_- quadrature variance. We obtain that, at 96% probability, $\langle P_-^2 \rangle < 0.5$ for the data in Fig. 3 (dataset A). This approach also yields the most likely value $\langle P_-^2 \rangle \approx 0.42 \pm 0.08$, which agrees with the values presented above but is determined independently.

The best estimate of the Duan quantity is found by combining all this information—namely, $\langle X_+^2 \rangle$, obtained from probe detection, and $\langle P_-^2 \rangle$, determined as explained above—which gives $\langle X_+^2 \rangle + \langle P_-^2 \rangle = 0.83 \pm 0.13$ for dataset A and $\langle X_+^2 \rangle + \langle P_-^2 \rangle = 0.72 \pm 0.18$ for dataset B, where the errors represent the worst-case combination of the individual measurements. The fluctuations hence satisfy the Duan bound for entanglement, $\langle X_+^2 \rangle + \langle P_-^2 \rangle < 1$, with confidence better than 2 standard deviations.

Entangled mechanical oscillators combined with phase-sensitive measurement systems can find practical use in the precise reconstruction of classical resonant forces, which has implications for quantum metrology. The entanglement of massive mechanical oscillators establishes a new regime for experimental quantum mechanics. In the future one could demonstrate quantum teleportation of motional states or, if phonon number measurements are possible, test Clauser–Horne–Shimony–Holt-type Bell inequalities[29] with massive mechanical objects.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The source data and codes for the main-text figures are available at http://doi.org/10.5281/zenodo.1205319.

Online content

1. Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777–780 (1935).
2. Aspect, A., Dalibard, J. & Roger, G. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.* **49**, 1804–1807 (1982).
3. Heidmann, A. et al. Observation of quantum noise reduction on twin laser beams. *Phys. Rev. Lett.* **59**, 2555–2557 (1987).
4. Ou, Z. Y., Pereira, S. F., Kimble, H. J. & Peng, K. C. Realization of the Einstein–Podolsky–Rosen paradox for continuous variables. *Phys. Rev. Lett.* **68**, 3663–3666 (1992).
5. Bowen, W. P., Treps, N., Schnabel, R. & Lam, P. K. Experimental demonstration of continuous variable polarization entanglement. *Phys. Rev. Lett.* **89**, 253601 (2002).
6. Jost, J. D. et al. Entangled mechanical oscillators. *Nature* **459**, 683–685 (2009).
7. Hensen, B. et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
8. Steffen, M. et al. Measurement of the entanglement of two superconducting qubits via state tomography. *Science* **313**, 1423–1425 (2006).
9. DiCarlo, L. et al. Preparation and measurement of three-qubit entanglement in a superconducting circuit. *Nature* **467**, 574–578 (2010).
10. Palomaki, T. A., Teufel, J. D., Simmonds, R. W. & Lehnert, K. W. Entangling mechanical motion with microwave fields. *Science* **342**, 710–713 (2013).
11. Leggett, A. J. Macroscopic quantum systems and the quantum theory of measurement. *Prog. Theor. Phys. Suppl.* **69**, 80–100 (1980).
12. Julsgaard, B., Kozhekin, A. & Polzik, E. S. Experimental long-lived entanglement of two macroscopic objects. *Nature* **413**, 400–403 (2001).
13. Lee, K. C. et al. Entangling macroscopic diamonds at room temperature. *Science* **334**, 1253–1256 (2011).
14. Klimov, P. V., Falk, A. L., Christle, D. J., Dobrovitski, V. V. & Awschalom, D. D. Quantum entanglement at ambient conditions in a macroscopic solid-state spin ensemble. *Sci. Adv.* **1**, e1501015 (2015).
15. Woolley, M. J. & Clerk, A. A. Two-mode squeezed states in cavity optomechanics via engineering of a single reservoir. *Phys. Rev. A* **89**, 063805 (2014).
16. Woolley, M. J. & Clerk, A. A. Two-mode back-action-evading measurements in cavity optomechanics. *Phys. Rev. A* **87**, 063846 (2013).
17. Ockeloen-Korppi, C. F. et al. Quantum backaction evading measurement of collective mechanical modes. *Phys. Rev. Lett.* **117**, 140401 (2016).
18. Mancini, S., Giovannetti, V., Vitali, D. & Tombesi, P. Entangling macroscopic oscillators exploiting radiation pressure. *Phys. Rev. Lett.* **88**, 120401 (2002).
19. Pinard, M. et al. Entangling movable mirrors in a double-cavity system. *Europhys. Lett.* **72**, 747–753 (2005).
20. Tan, H., Li, G. & Meystre, P. Dissipation-driven two-mode mechanical squeezed states in optomechanical systems. *Phys. Rev. A* **87**, 033829 (2013).
21. Wang, Y.-D. & Clerk, A. A. Reservoir-engineered entanglement in optomechanical systems. *Phys. Rev. Lett.* **110**, 253601 (2013).
22. Li, J., Haghighi, I. M., Malossi, N., Zippilli, S. & Vitali, D. Generation and detection of large and robust entanglement between two different mechanical resonators in cavity optomechanics. *New J. Phys.* **17**, 103037 (2015).
23. Mahboob, I., Okamoto, H., Onomitsu, K. & Yamaguchi, H. Two-mode thermal-noise squeezing in an electromechanical resonator. *Phys. Rev. Lett.* **113**, 167203 (2014).
24. Pontin, A. et al. Dynamical two-mode squeezing of thermal fluctuations in a cavity optomechanical system. *Phys. Rev. Lett.* **116**, 103601 (2016).
25. Wollman, E. E. et al. Quantum squeezing of motion in a mechanical resonator. *Science* **349**, 952–955 (2015).
26. Pirkkalainen, J.-M., Damskägg, E., Brandt, M., Massel, F. & Sillanpää, M. A. Squeezing of quantum noise of motion in a micromechanical resonator. *Phys. Rev. Lett.* **115**, 243601 (2015).
27. Lecocq, F., Clark, J. B., Simmonds, R. W., Aumentado, J. & Teufel, J. D. Quantum nondemolition measurement of a nonclassical state of a massive object. *Phys. Rev. X* **5**, 041037 (2015).
28. Duan, L.-M., Giedke, G., Cirac, J. I. & Zoller, P. Inseparability criterion for continuous variable systems. *Phys. Rev. Lett.* **84**, 2722–2725 (2000).
29. Banaszek, K. & Wódkiewicz, K. Nonlocality of the Einstein–Podolsky–Rosen state in the Wigner representation. *Phys. Rev. A* **58**, 4345–4347 (1998).
30. Weedbrook, C. et al. Gaussian quantum information. *Rev. Mod. Phys.* **84**, 621–669 (2012).
31. Teufel, J. D. et al. Sideband cooling of micromechanical motion to the quantum ground state. *Nature* **475**, 359–363 (2011).

Author contributions M.A.S. initiated the project and was involved in all subsequent stages. C.F.O.-K. carried out the measurements. C.F.O.-K. and E.D. analysed the data. E.D. and J.-M.P. designed and fabricated the devices. M.J.W., A.A.C., F.M. and M.A. developed the theory. All authors participated in the writing of the paper.

Competing interests The authors declare no competing interests.

Additional information

LETTER

Low–loss plasmon–assisted electro–optic modulator

Christian Haffner[1]*, Daniel Chelladurai[1], Yuriy Fedoryshyn[1], Arne Josten[1], Benedikt Baeuerle[1], Wolfgang Heni[1], Tatsuhiko Watanabe[1], Tong Cui[1], Bojun Cheng[1], Soham Saha[3], Delwin L. Elder[2], Larry. R. Dalton[2], Alexandra Boltasseva[3], Vladimir M. Shalaev[3], Nathaniel Kinsey[4] & Juerg Leuthold[1]*

For nearly two decades, researchers in the field of plasmonics[1]—which studies the coupling of electromagnetic waves to the motion of free electrons near the surface of a metal[2]—have sought to realize subwavelength optical devices for information technology[3–6], sensing[7,8], nonlinear optics[9,10], optical nanotweezers[11] and biomedical applications[12]. However, the electron motion generates heat through ohmic losses. Although this heat is desirable for some applications such as photo-thermal therapy, it is a disadvantage in plasmonic devices for sensing and information technology[13] and has led to a widespread view that plasmonics is too lossy to be practical. Here we demonstrate that the ohmic losses can be bypassed by using 'resonant switching'. In the proposed approach, light is coupled to the lossy surface plasmon polaritons only in the device's off state (in resonance) in which attenuation is desired, to ensure large extinction ratios between the on and off states and allow subpicosecond switching. In the on state (out of resonance), destructive interference prevents the light from coupling to the lossy plasmonic section of a device. To validate the approach, we fabricated a plasmonic electro-optic ring modulator. The experiments confirm that low on-chip optical losses, operation at over 100 gigahertz, good energy efficiency, low thermal drift and a compact footprint can be combined in a single device. Our result illustrates that plasmonics has the potential to enable fast, compact on-chip sensing and communications technologies.

Telecommunication devices such as electro-optic modulators must feature low insertion loss (the reduction of signal power resulting from insertion of the device into the transmission line) while providing a large change in phase (Δn) or amplitude ($\Delta \alpha$) accumulated over a short device length[14]. Beyond that, modulators should offer low driving voltages and high-speed operation.

In recent years, silicon photonics has emerged that has low propagation losses (of the order of dB cm^{-1}) but has struggled to achieve large modulation depth (that is, Δn and $\Delta \alpha$) for submillimetre devices[15]. State-of-the-art devices maximize their modulation through the use of resonant structures[16], enabling compact (few square micrometres) and energy-efficient components[17,18]. And although large quality factors (Q-factors, large values indicate low energy losses), of the order of several thousand, allow considerable reduction of the driving voltage, they limit the speed and increase the sensitivity of devices to temperature and fabrication fluctuations[19]. More recently, surface plasmonic polariton (SPP) devices have exploited the extreme confinement of light to achieve exceptional modulation of phase and amplitude within a few micrometres[3,20]. However, the metals that give plasmonics such promise are also the largest hindrance, as such devices suffer from large on-state loss (of the order of dB µm^{-1}). To combat plasmonic losses, some devices use hybrid plasmonic–photonic modes[4,6,21] whereas others minimize the length of the active section[5,22]. Still, typical insertion losses of 10 dB due to plasmonic propagation and photonic-to-plasmonic mode conversion remain a concern for high-speed state-of-the-art devices[5,6]. Thus, plasmonic on-chip technologies have been unable to replace the existing photonic or electronic solutions.

We propose an approach in which losses in plasmonic waveguides can be selectively used or bypassed to achieve low insertion loss, strong modulation, compact footprint and high speed, simultaneously. Previously, the goal has been to reduce the ohmic loss as much as possible to minimize the device's loss in the on state. Instead, we show that plasmonic losses can be tolerated by designing the device geometry such that light passes through the lossy section when required (in the off state). To achieve this, we use a plasmonic ring resonator coupled to a buried low-loss silicon photonic waveguide (see Fig. 1).
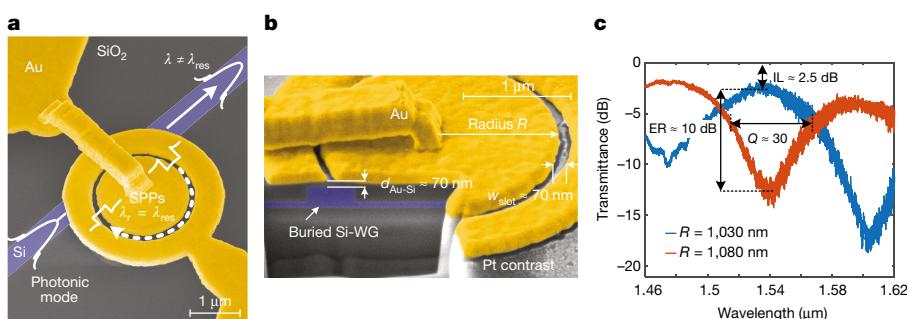
Fig. 1 | False-coloured scanning electron microscopy (SEM) image of a plasmonic ring resonator and the corresponding transmittance. a, Top view; **b**, cross-section of the resonator. Photonic modes propagating in the buried silicon waveguide (Si-WG) resonator couple partially to the SPPs in the MIM ring when the resonance condition is fulfilled. While out of resonance, operation results in low-loss light transmission. $d_{\text{Au–Si}}$, vertical distance between the gold ring resonator and the buried silicon. **c**, Passive measurements of two identical ring resonators that differ only in radii (blue, 1,030 nm; red, 1,080 nm). To estimate the optical properties of such plasmonic ring resonators through passive measurement, we compare the two resonators to determine the insertion losses (IL), extinction ratio (ER) and Q-factor at 1.54 µm. Here we use the off-resonance resonator (blue) to measure the insertion loss, and the on-resonance resonator (red) to extract the Q-factor at 1.54 µm. The difference between the red and blue reveals the extinction ratio. With the resonant approach, insertion losses of about 2.5 dB are measured with extinction ratios above 10 dB.

[1]ETH Zurich, Institute of Electromagnetic Fields (IEF), Zurich, Switzerland. [2]Department of Chemistry, University of Washington, Seattle, WA, USA. [3]School of Electrical & Computer Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN, USA. [4]Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, VA, USA. *e-mail: haffnerc@ethz.ch; leuthold@ethz.ch
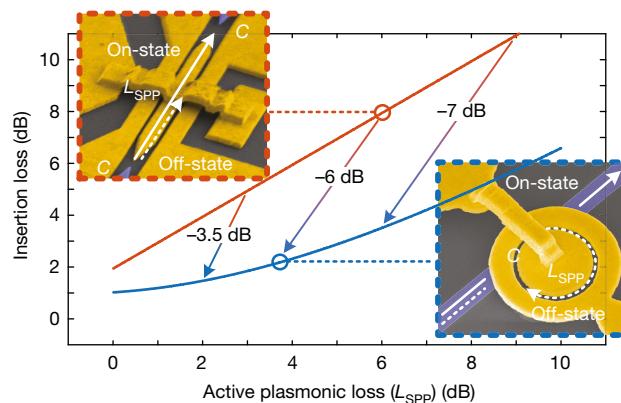
Fig. 2 | Theoretical loss advantage of critical coupled resonant device over non-resonant push–pull MZ device. The insertion losses are plotted against the active plasmonic loss (L_{SPP}) in the slot waveguide of the MZ modulator (left inset, red) and resonator (right inset, blue). Losses can be reduced by more than 6 dB. This is for the following reasons: (i) the bypassing mechanism: only a fraction of light experiences plasmonic losses; (ii) resonant enhancement: the resonators achieve the same modulation depth (ΔT) for shorter active plasmonic sections than their non-resonant counterparts (indicated by the horizontal shift of the arrows); and (iii) coupling scheme: non-resonant approaches require two photonic/SPP converters whereas resonant approaches require only one.

Unlike previous approaches, which aim to reduce the resonator loss to realize high-Q cavities or lasing[7,23–26], our resonator exhibits typical propagation losses within the plasmonic cavity and mostly bypasses this lossy section in the on state through destructive interference. Using this design, we demonstrate a plasmonic modulator that can meet the key performance metrics of modern optical communications links.

Figure 1 shows the proposed device geometry, which comprises a gold metal–insulator–metal (MIM) slot waveguide ring coupled to a buried silicon bus waveguide, forming a notch filter with a resonant wavelength of λ_{res} (see section II of the Supplementary Information). The slot waveguide is filled with an organic electro-optic (OEO) material which alters the device's resonance condition through the Pockels effect (Δn_{SPP})[27]. This permits fast and selective use of the resonator's ohmic loss to attenuate the signal ($\Delta \alpha$) in the bus waveguide by applying a voltage.

Figure 1c shows the measured transmittance over the wavelength of two representative devices that differ slightly in radius. The difference in radius does not affect their optical properties (insertion loss, extinction ratio, Q-factor), but results in a distinct off-resonance ($\lambda_0 \neq \lambda_{res}$) and on-resonance ($\lambda_0 = \lambda_{res}$) behaviour at the telecommunication wavelength of 1.54 μm. The off-resonance ring reveals an insertion loss of about 2.5 dB, the on-resonance ring shows a Q-factor of about 30, and their difference in transmittance indicates an extinction ratio of about 10 dB. We note that non-resonant devices based on

similar-length MIM waveguides feature an insertion loss ranging from 8 dB to 10 dB (refs. [5,6,22]).

The reduced insertion loss can be understood by comparing the exemplary operating principles of a non-resonant Mach–Zehnder modulator (MZ)[5] (Fig. 2, left inset; white solid arrow indicates the plasmonic section) and a resonant ring (Fig. 2, right inset). The insertion loss of the device is a function of its coupling efficiency, geometry and accumulated ohmic loss. In both concepts, light couples to and from the plasmonic structure with a coupling efficiency C. A transmission modulation is then induced by the Pockels effect over an active plasmonic section of length or circumference l. The modulation depth ΔT, where T is transmittance, and also the loss L_{SPP} scale with the length of the active plasmonic section[6]. Figure 2 shows the overall insertion loss over the L_{SPP} in the active plasmonic area for a MZ (red curve) and for a critically coupled resonator (blue curve). The arrows in the plot indicate the performance of devices with an equal ΔT.

The following points arise from Fig. 2. First, the resonator's total loss is always smaller than that of a MZ modulator (blue curve < red curve), owing to the bypassing mechanism[28]. The insertion loss increases with L_{SPP} (that is, inverse to the cavity's Q-factor) as the critical coupling condition requires that more light couples to the plasmonic section. Second, the resonator has an insertion loss 1 dB lower than that of the MZ at $L_{SPP} \approx 0$ dB. This is because the non-resonant device requires two photonic–SPP converters as both the on state and the off state propagate through the plasmonic section. Ohmic losses in the converters limit the conversion efficiency to about 1 dB (refs. [22,29]). In contrast, the selection mechanism of the resonator (on state: bus waveguide; off state: ring) requires a converter which couples only a fraction of the light to the ring. Third, in the ring we take advantage of the resonantly enhanced ΔT to reduce the device length (see section VII of the Supplementary Information). For our resonant structures, we have calculated an enhancement of about 1.5. Consequently, our ring with a circumference of $l = 6$ μm ($L_{SPP} \approx 4$ dB) offers the same transmittance change as a MZ of $l = 9$ μm ($L_{SPP} \approx 6$ dB). In total, the ring device offers an advantage of 6 dB insertion loss over the MZ modulator. Losses can be further reduced by under-coupling the resonator, as limited extinction ratios of 10 dB are sufficient for many practical applications[30].

To illustrate the modulation performance (ΔT) of the plasmonic resonator, the effective refractive index for SPPs (Δn_{SPP}) is altered by applying a bias between the inner and outer rings[27]. Figure 3a shows the transmitted power versus applied voltage for a probe wavelength of approximately 1.52 μm. We observed an insertion loss < 3 dB, an extinction ratio of about 10 dB and a linear response (dashed green line) for a peak voltage of 3.5 V with an extinction ratio of about 6 dB. This performance in terms of insertion loss and extinction ratio is similar to that of well-developed complementary metal oxide semiconductor (CMOS) photonic resonators[19]. We estimate that operation under a digital peak-to-peak driving voltage (1 V)[19] will soon be achievable by using other plasmonic materials such as silver or copper, using the newest OEO materials and improving fabrication (see Methods: Future
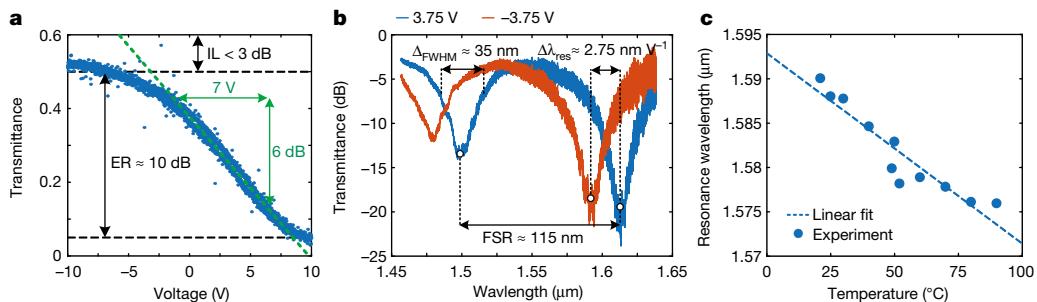
Fig. 3 | Sensitivity and stability of the plasmonic resonator. a, Voltage sensitivity of the resonator's transmittance. IL, insertion loss; ER, extinction ratio. Probe wavelength 1.52 μm. **b,** Sensitivity of the ring as a function of the wavelength. A change ($\Delta n_{slot} \approx 0.03$) in the refractive index of the slot-filling material causes a large change in the resonance

wavelength (blue/red). FSR, free spectral range. **c,** The resonator shows stable operation across a large thermal variation. The dashed line indicates a slope of $\Delta \lambda_{res}/(\Delta_{FWHM}\Delta T) \approx 0.4\%$ K^{-1}. These characteristics make the plasmonic MIM-ring resonator a promising candidate in the field of optical modulators and sensors.
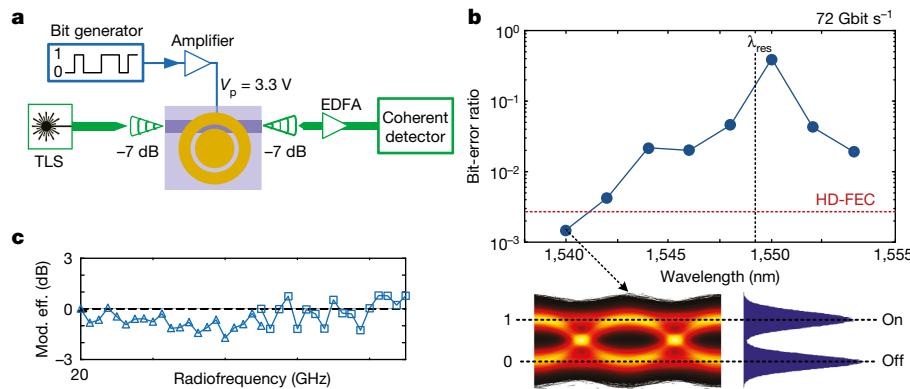
Fig. 4 | High-speed data experiments with a plasmonic ring resonator used as an electro-optic modulator. a, The experimental set-up. EDFA, erbium-doped fibre amplifier; TLS, tuneable laser source; V_p, peak voltage. **b**, BER versus wavelength for a resonator with $\lambda_{res} = 1{,}549$ nm. BERs below the HD-FEC limit show successful data modulation and detection without the use of a temperature control. The BER increases at the resonance wavelength as expected from the notch-filter response of the resonator. **c**, The bandwidth of the plasmonic resonator, which is beyond 110 GHz. Mod. eff., modulation efficiency.

improvements). To highlight the mechanism of the modulation, the transmission of the device was measured under a positive and negative bias (± 3.75 V) as a function of wavelength (see Fig. 3b). This results in a normalized sensitivity, $S_\lambda / \Delta_{FWHM} = \Delta\lambda_{res}/(\Delta_{FWHM}\Delta n)$, where Δ_{FWHM} is the full-width at half-maximum, of about 17 RIU^{-1} (RIU is refractive index unit), assuming a relative change in the OEO refractive index of about 0.03. This can be compared to commercial SPR sensors that achieve values of 50 RIU^{-1} by using a free-space Kretschmann configuration[31]. Therefore, our approach is already close to the non-integrated free-space approaches.

The moderate Q-factor guarantees a high operational speed and provides good thermal stability. For example, the measured resonance frequency is plotted in Fig. 3c over a temperature range from 20 °C up to 90 °C. In this case, the resonant frequency is found to follow a linear trend line with a slope of about 0.4% K^{-1}. The insensitivity of the low-Q plasmonic resonator to temperature fluctuations is in strong contrast to high-Q photonic resonators, which are two orders of magnitude more sensitive to temperature fluctuations (approximately 100% K^{-1})[18]. This is extremely beneficial for applications in which strong temperature fluctuations occur. For example, a low-Q plasmonic resonant sensor or modulator would be immune to thermal fluctuations of ± 5 °C which normally occur in CPUs, whereas high-Q photonic resonators require power-consuming temperature controls to maintain operation[19]. Furthermore, the moderate Q factors are beneficial for high-speed operation, as desired in electro-optic modulators. As a result, we are able to push the bandwidth of a resonant electro-optic modulator well beyond 100 GHz (see Methods section 'Speed limitations'). In comparison, photonic resonators are more likely to be limited to bandwidths of about 20 GHz and below[18,19].

Subsequently, we performed high-speed data experiments to demonstrate the robustness, high speed and low power switching capability. Although the 'resonant switching' principle can also be used to optimize sensors, we focus on high-speed applications because their sensitivity and stability requirements are stricter.

In the experiment of Fig. 4a, the peak driving voltage was approximately 3.3 V, and the laser wavelength was varied to capture the response of the modulator. Low device losses of 2.5 dB and fibre-to-silicon waveguide losses of about 7 dB resulted in fibre-to-fibre coupling losses of 16.5 dB, enabling successful operation with low laser powers of 4 dBm and below. The resulting bit-error ratio (BER) versus wavelength for a 72 Gbit s^{-1} signal is shown in Fig. 4b, where a peak in the BER is observed at λ_{res} (dashed line). The BER shows a peak because the change of transmittance (ΔT) has a symmetric response to voltage when the probing wavelength λ equals λ_{res}. Consequently, we confirmed that the operating mechanism relies on amplitude modulation. Off-resonance, the BER quickly dropped to about 1×10^{-3}, which is

below the hard-decision forward error correction (HD-FEC) limit and allows for successful data modulation and detection[32]. No thermal heater was required for stabilization. Additionally, we reduced the data rate to 36 Gbit s^{-1} and 18 Gbit s^{-1} and found BERs of about 2×10^{-6} and $< 1 \times 10^{-6}$, respectively, indicating that the BER at 72 Gbit s^{-1} is mainly limited by the electrical equipment (see Methods: Data experiments). We estimate the energy consumption of the modulator to be about 12 fJ bit^{-1} at 72 Gbit s^{-1} (ref. [14]) for a device capacitance of 1.1 fF.

We have demonstrated that low-Q resonant designs using highly confined SPPs can enable low-loss active plasmonic devices with a good modulation depth. We believe that our approach—unlike conventional resonant photonics—breaks the trade-off between sensitivity (high-Q) on the one hand and speed and temperature stability (low-Q) on the other. Our work can be seen as a step towards practical plasmonics that ultimately serves as a compact and fast gateway between electronics (local signal processing) and photonics (broad bandwidth and low-loss data stream). The proposed slot waveguide approach could be of interest for sensing, because the resonant response could be exploited for many other material systems ranging from low-index materials such as aqueous solutions to high-index materials such as silicon.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0031-4.

1. Maier, S. A. et al. Plasmonics—a route to nanoscale optical devices. *Adv. Mater.* **13**, 1501–1505 (2001).
2. Gramotnev, D. K. & Bozhevolnyi, S. I. Plasmonics beyond the diffraction limit. *Nat. Photonics* **4**, 83–91 (2010).
3. Dionne, J. A., Diest, K., Sweatlock, L. A. & Atwater, H. A. PlasMOStor: a metal−oxide−Si field effect plasmonic modulator. *Nano Lett.* **9**, 897–902 (2009).
4. Sorger, V. J., Lanzilotti-Kimura Norberto, D., Ma, R.-M. & Zhang, X. Ultra-compact silicon nanophotonic modulator with broadband response. *Nanophotonics* **1**, 17–22 (2012).
5. Haffner, C. et al. All-plasmonic Mach–Zehnder modulator enabling optical high-speed communication at the microscale. *Nat. Photon.* **9**, 525–528 (2015).
6. Keeler, G. A. et al. in *Optical Fiber Communication Conference* Th3I.1 (Optical Society of America, 2017).
7. Min, B. et al. High-Q surface-plasmon-polariton whispering-gallery microcavity. *Nature* **457**, 455–458 (2009).
8. Brolo, A. G. Plasmonics for future biosensors. *Nat. Photonics* **6**, 709–713 (2012).
9. Cai, W., Vasudev, A. P. & Brongersma, M. L. Electrically controlled nonlinear generation of light with plasmonics. *Science* **333**, 1720–1723 (2011).
10. Kauranen, M. & Zayats, A. V. Nonlinear plasmonics. *Nat. Photonics* **6**, 737–748 (2012).

11. Ndukaife, J. C. et al. Long-range and rapid transport of individual nano-objects by a hybrid electrothermoplasmonic nanotweezer. *Nat. Nanotech.* **11**, 53–59 (2016).
12. Hirsch, L. R. et al. Nanoshell-mediated near-infrared thermal therapy of tumors under magnetic resonance guidance. *Proc. Natl Acad. Sci. USA* **100**, 13549–13554 (2003).
13. Khurgin, J. B. How to deal with the loss in plasmonics and metamaterials. *Nat. Nanotech.* **10**, 2–6 (2015).
14. Miller, D. A. B. Attojoule optoelectronics for low-energy information processing and communications; a tutorial review. *J. Lightwave Technol.* https://doi.org/10.1109/JLT.2017.2647779 (2017).
15. Reed, G. T., Mashanovich, G., Gardes, F. Y. & Thomson, D. J. Silicon optical modulators. *Nat. Photon.* **4**, 518–526 (2010); erratum *Nat. Photon.* **4**, 660 (2010).
16. Vahala, K. J. Optical microcavities. *Nature* **424**, 839–846 (2003).
17. Xu, Q., Schmidt, B., Pradhan, S. & Lipson, M. Micrometre-scale silicon electro-optic modulator. *Nature* **435**, 325–327 (2005).
18. Timurdogan, E. et al. An ultralow power athermal silicon modulator. *Nat. Commun.* **5**, 4008 (2014).
19. Sun, C. et al. Single-chip microprocessor that communicates directly using light. *Nature* **528**, 534–538 (2015).
20. Emboras, A. et al. Atomic scale plasmonic switch. *Nano Lett.* https://doi.org/10.1021/acs.nanolett.5b04537 (2015).
21. Oulton, R. F., Sorger, V. J., Genov, D. A., Pile, D. F. P. & Zhang, X. A hybrid plasmonic waveguide for subwavelength confinement and long-range propagation. *Nat. Photon.* **2**, 496–500 (2008).
22. Zhu, S., Lo, G. Q. & Kwong, D. L. Phase modulation in horizontal metal–insulator–silicon–insulator–metal plasmonic waveguides. *Opt. Express* **21**, 8320–8330 (2013).
23. Hill, M. T. et al. Lasing in metallic-coated nanocavities. *Nat. Photon.* **1**, 589–594 (2007).
24. Bozhevolnyi, S. I., Volkov, V. S., Devaux, E., Laluet, J.-Y. & Ebbesen, T. W. Channel plasmon subwavelength waveguide components including interferometers and ring resonators. *Nature* **440**, 508–511 (2006).
25. Kress, S. J. P. et al. A customizable class of colloidal-quantum-dot spasers and plasmonic amplifiers. *Sci. Adv.* **3**, e1700688 (2017).
26. Zhu, W. et al. Surface plasmon polariton laser based on a metallic trench Fabry–Perot resonator. *Sci. Adv.* **3**, e1700909 (2017).
27. Sun, S. S. & Dalton, L. R. *Introduction to Organic Electronic and Optoelectronic Materials and Devices* 28–37 (Taylor & Francis, Boca Raton, 2008).
28. Zanotto, S., Morichetti, F. & Melloni, A. Fundamental limits on the losses of phase and amplitude optical actuators. *Laser Photonics Rev.* **9**, 666–673 (2015).
29. Delacour, C. et al. Efficient directional coupling between silicon and copper plasmonic nanoslot waveguides: toward metal−oxide−silicon nanophotonics. *Nano Lett.* **10**, 2922–2926 (2010).
30. Bogaerts, W. et al. Silicon microring resonators. *Laser Photonics Rev.* **6**, 47–73 (2012).
31. Becker, J., Trügler, A., Jakab, A., Hohenester, U. & Sönnichsen, C. The optimal aspect ratio of gold nanorods for plasmonic bio-sensing. *Plasmonics* **5**, 161–167 (2010).
32. Chang, F., Onohara, K. & Mizuochi, T. Forward error correction for 100 G transport networks. *IEEE Commun. Mag.* **48**, https://doi.org/10.1109/MCOM.2010.5434378 (2010).

Author contributions C.H., N.K., V.M.S., A.B. and J.L. conceived the concept and supervised the project. C.H., D.C., S.S. and T.C. designed the modulator and developed the analytical framework for fast optimization. T.W. designed the photonic grating coupler. C.H., D.C. and Y.F. fabricated the modulator and developed the required process technology. B.C. developed a focused ion beam process to image the cross-section with minimal destructive influence on the suspended bridge. D.L.E., W.H., C.H. and L.R.D. developed, synthesized and implemented the poling procedure of the OEO material for plasmonic ring resonators. C.H. and J.L. designed the experiments. C.H., D.C. and T.C. performed the passive characterization. C.H. performed the temperature sensitivity, d.c. switching and electro-optic bandwidth experiments. B.B., A.J. and C.H. performed the high-speed data experiment. B.B. and A.J. designed, calibrated and automated the high-speed data experiment. B.B. and A.J. developed the digital-signal processing for data generation and analysis of the high-speed data experiment. All authors discussed and analysed the data. C.H., N.K., D.C. and J.L. wrote the manuscript.

Competing interests The authors declare no competing financial interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0031-4.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0031-4.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to C.H. or J.L.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Future improvements. The plasmonic material strongly influences the performance of the ring resonator in terms of losses and resonant enhancement. In this section, we discuss the influence of various plasmonic materials and we show that the modulator performance improves when replacing gold with copper (CMOS-compatible) or silver (lowest loss).

In a first step, we investigate the unloaded Q-factor for different plasmonic metals. Extended Data Fig. 1 shows the unloaded quality factor Q as a function of the slot width and the radius for Au (panel a), Cu (b) and Ag (c).

Both Cu and Ag feature higher Q-factors than Au. When changing the plasmonic metal from Au to Ag, the Q-factor is more than doubled for geometries dominated by propagation loss (140 for Ag versus 60 for Au). Each material shows the trade-off between propagation loss and bending loss resulting in the characteristic triangular contour lines (see section III of the Supplementary Information). Furthermore, using low-loss materials such as Ag will increase not only the intrinsic Q-factor but also the coupling efficiency C. Three-dimensional simulations indicate that C increases from 0.8 to 0.9.

Besides the plasmonic material, the OEO material provides an additional path for further improvement. Recent experiments have shown that electro-optic coefficients over $300\,\mathrm{pm\,V^{-1}}$ are achievable by operating in close vicinity of the OEO material resonance ($\lambda \approx 1{,}250\,\mathrm{nm}$)[33]. Additionally, other OEO materials such as JRD1 show electro-optic coefficients over $300\,\mathrm{pm\,V^{-1}}$ (refs. [34–36]).

To evaluate the influence of these improvements on the modulator performance, Extended Data Fig. 1 shows the wavelength-dependent transmittance of a MIM-ring resonator obtained by the analytic model (equation (2) in the Supplementary Information). Extended Data Fig. 1d shows the transmittance for Au while Extended Data Fig. 1e includes the possible improvements with Ag and JRD1 (electro-optic coefficient of $300\,\mathrm{pm\,V^{-1}}$). The radius is $1\,\mu\mathrm{m}$, the slot width is $80\,\mathrm{nm}$ and the outer electrode height is $350\,\mathrm{nm}$. The solid lines correspond to positive (blue) and negative (red) biasing of 2 V.

Such improved configurations offer through-port insertion loss as small as 0.5 dB (Ag) compared with 1.5 dB (Au). Furthermore, the switching sensitivity is increased because of the higher nonlinearity (sensitivity increase by $3\times$) and the larger Q-factor ($2\times$). Thus, the absolute insertion loss reduces from 4.3 dB to 1.2 dB, and the extinction ratio increases from 2 dB to 18 dB when driven with $\pm 2\,\mathrm{V}$ at the chosen probing wavelength (vertical green line, Extended Data Fig. 1d, e). Thus, low-power CMOS transistors may drive the plasmonic resonators without the help of electrical amplification. This would reduce the total energy consumption of the modulators to below 1 fJ per bit.

Alternative active materials for modulation and sensing. Here, we discuss the influence of the refractive index of the slot material on the unloaded Q-factor of plasmonic ring resonators. The active material could be either an OEO material for modulation or a biosensitive material for sensing purposes. Extended Data Fig. 2 shows the simulated Q-factors for different slot widths and radii. The simulated refractive indices range from 1.33 up to 3.48 as shown in subplots in Extended Data Fig. 2a to h.

All material systems achieve Q-factors above 30. The refractive index of the active material strongly influences the bending and propagation losses. In the case of low n_{slot} for Extended Data Fig. 2a and b, the Q-factor is mainly limited by bending losses for all slot widths as indicated by the diagonal contour lines (dashed lines). This behaviour changes when n_{slot} increases to 1.75 (Extended Data Figs. 2c) or 2 (Extended Data Fig. 2d) and typical triangular contour lines can be observed highlighting the trade-off between bending losses and propagation losses (see section IV of the Supplementary Information). A further increase of the refractive index results in dominating propagation losses. The triangular shape of the contour line changes to contour lines parallel to the bottom axis (radius). In this region (see subplots in Extended Data Fig. 2e to h), the Q-factor drops for the same geometrical parameters as propagation loss increases with the refractive index[37].

The subplots shown in Extended Data Fig. 2a, c, e and h are of most interest as they represent materials that can be used for sensing: (a) $n_{\mathrm{slot}} = 1.33$ ($\mathrm{H_2O}$); or for electro-optic modulation: (c) $n_{\mathrm{slot}} = 1.75$ (OEO)[35]; (e) $n_{\mathrm{slot}} = 2.25$ (barium titanate, BTO)[38]; and (h) $n_{\mathrm{slot}} = 3.48$ (Si)[22].

Sample fabrication. All reported devices were processed on silicon-on-insulator wafers with a 220-nm-thick Si device layer and 3-μm-thick buried oxide. Electron beam lithography (Vistec EBPG5200) in combination with ICP-RIE (Oxford Plasmalab System 100) were applied to pattern the Si photonic components. A positive tone resist (poly(methyl methacrylate), PMMA) and a negative tone resist (hydrogen silsesquioxane, HSQ) were used as masks to define the partially (70 nm) and fully (220 nm) etched Si structures, respectively[39]. Planarization of the patterned Si structures was realized by spin-coating the samples with a 1-μm-thick HSQ layer followed by rapid thermal annealing to convert HSQ to $\mathrm{SiO_2}$ (ref. [40]). We observed thickness variations of $\pm 10\,\mathrm{nm}$ with a periodicity of $50\,\mu\mathrm{m}$ throughout the chip. This does not affect device performance, as the footprint of the resonators is much smaller than the periodicity. We controlled the coupling between the ring and bus waveguide by using the following protocol, which enabled us to adjust the average vertical spacing between the gold ring and silicon bus ($d_{\mathrm{Au-Si}}$) to within 5 nm of the target value. First, we use an electron-beam-patterned PMMA resist mask to open silicon etch marks placed next to the device to measure the HSQ thickness during the following wet etching process. Afterwards, we used a second e-beam patterned PMMA resist to locally etch down the $\mathrm{SiO_2}$ over the bus waveguide to the desired spacing $d_{\mathrm{Au,Si}} = 70 \pm 10\,\mathrm{nm}$ by using a buffered HF (7:1) solution. The slow etch-rate of $1.2\,\mathrm{nm\,s^{-1}}$ enables good process control. Note that the spin-coating process rather than the wet etching process introduces the uncertainty of $\pm 10\,\mathrm{nm}$. Contact pads, outer rings (height $h = 350\,\mathrm{nm}$), inner rings ($h = 150\,\mathrm{nm}$) and bridges ($h = 350\,\mathrm{nm}$) were produced with a sequence of e-beam lithography, e-beam evaporation of Au and lift-off processes as reported in ref. [5]. Extended Data Fig. 3 shows a tilted top view of the fabricated ring resonator before the bridge fabrication step. The height of the outer electrode is 350 nm, whereas the inner electrode is only 150 nm, enabling a reduction of the bending losses (see section IV of the Supplementary Information).

We characterized the geometry of all devices using SEM for better comparison to simulations. This allows us to estimate parameters such as the electro-optic coefficient, r_{33}. Finally, the binary OEO material (75%HD-BB-OH/25%YLD124)[41] was spin-coated then poled at its glass transition temperature (115 °C) by applying electric fields of approximately $130\,\mathrm{V\,\mu m^{-1}}$. It has been shown[33] that this OEO material allows operation up to 80 °C.

We chose the following design parameter sets: slot widths of 80 nm and 100 nm with radii ranging from 900 nm to 1,200 m and 1,400 nm to 1600 nm.

Passive characterization. We determined the transmittance of the resonators by subtracting back-to-back measurements of structures that comprise the same grating couplers and photonic waveguides but lack the plasmonic section. The standard deviation of the back-to-back measurements was below $\pm 0.5\,\mathrm{dB}$ over the whole wavelength range of the tunable laser (1,465 nm to 1,635 nm). We measured the passive spectrum for various input power up to 10 dBm and did not observe any drift of the resonance wavelength due to self-heating. To prevent any self-heating, we set the laser power to 0 dBm for the passive characterization.

Extended Data Fig. 4 shows two histograms of the measured insertion loss (Extended Data Fig. 4a) and extinction ratio (Extended Data Fig. 4b) of 23 ring resonators with targeted slot widths of 80 nm. The average insertion loss is about 2.5 dB and the mean extinction ratio is about 9.7 dB.

We designed rings with varying radii to account for fabrication variations. Extended Data Fig. 4c shows the passive spectrum of ring resonators with radii ranging from 940 nm up to 990 nm. The resonance wavelength shifts with increasing radius by $\delta\lambda_{\mathrm{res}}/\delta R \approx 1.1$.

Sensitivity experiments. We used the same optical set-up as the passive characterization measurements to estimate the d.c./low-frequency electro-optic response of six devices. In a first experiment, we applied a 100 Hz triangular shaped electrical signal with a peak voltage amplitude of 10 V. We measured the modulation of the optical signal when operating the resonator around its 3 dB point. In a second experiment, we applied a d.c. bias and measured the resonators' wavelength shift as a function of the voltage. Extended Data Table 1 provides an overview of the insertion loss, extinction ratio and sensitivity of the resonance wavelength ($\Delta\lambda_{\mathrm{res}}$ per volt) as extracted from these two experiments. The extinction ratio (6–10 dB) and insertion loss (2–3 dB) during active switching match the values obtained from passive characterization. This shows that the modulators can be driven from minimum to maximum output power before dielectric breakdown occurs. Three-dimensional simulation of device 79 revealed a $\Delta\lambda_{\mathrm{res}}/V \approx 3.1\,\mathrm{nm\,V^{-1}}$ for an electro-optic coefficient of $100\,\mathrm{pm\,V^{-1}}$. Thus, we estimate an in-device electro-optic coefficient of about $90\,\mathrm{pm\,V^{-1}}$ and a $\Delta n \approx 0.03$ when biasing with $\pm 3.75\,\mathrm{V}$ based on the Pockels effect. Device 79 (last row) is shown in Fig. 3a and b of the main text.

Speed limitations. The electro-optic bandwidth of resonant electro-optic modulators is often determined by the photon lifetime of the resonant cavity. In the case of a ring-based electro-optic modulator with a Q-factor of 6,000, this equates to an electro-optic bandwidth of a few tens of gigahertz[42]. In the case of resonant plasmonic structures, the photon lifetime does not limit the electro-optic bandwidth, as these structures feature Q-factors that are two orders of magnitude smaller than the Q-factors of photonic approaches.

In general, the slowest of the following effects limits the electro-optic bandwidth of an electro-optic-modulator: (i) the RC time constant; (ii) the timescale of the nonlinear effect (here the Pockels effect); and (iii) the cavity lifetime. In the case of the plasmonic MIM ring modulator presented in this work, these contributions can be approximated as follows:

• τ_{RC}: The calculated capacitance of the MIM ring is approximately 1.1 fF (electrostatic 3D simulation), and its resistance is a few ohms with the nanometre-scale bridge being the main contributor[5]. This results in an RC bandwidth much greater than 1 THz.

• $\tau_{Pockels}$: The Pockels effect is based on the reorganization of π-electrons within a nanometre-long molecule and is estimated to be on the timescale of femto- or attoseconds[27]. This results in a Pockels bandwidth much greater than 1 THz.

• τ_{cavity}: The SPP lifetime is given by the quality factor divided by the angular frequency ($\tau_{res} \approx Q/\omega_{photon}$). The resonant structures presented in this work feature Q-factors of 50, while future implementations may reach values of up to 150. Thus, the cavity lifetime of SPPs is about 100 fs or less, resulting in a cavity lifetime bandwidth limit of more than 1 THz.

Based on these approximations, we expect, at worst, a terahertz bandwidth for a plasmonic MIM resonator using the Pockels effect. To confirm that bandwidth limitations are not an issue for resonant plasmonic modulators we performed electro-optic bandwidth measurements with a plasmonic resonator of radius 1.5 μm and a designed slot width of 100 nm. These design parameters are chosen because they feature the largest Q-factors and therefore the longest cavity lifetimes. Extended Data Fig. 5a shows the wavelength-dependent transmittance of the plasmonic resonators and Extended Data Fig. 5b shows the normalized modulation efficiency as a function of the modulation frequency. The ratio between the unmodulated carrier and the modulated sidebands is the modulation efficiency which was measured with an optical spectrum analyser and normalized with respect to the lowest applied frequency[43].

The different colours correspond to three characteristic probing wavelengths. Red, yellow and purple are the off-resonance, 3 dB point and on-resonance probing wavelengths, respectively. For all three curves, a similar behaviour is observed, and electro-optic bandwidths beyond 100 GHz are found. This indicates that the cavity lifetime is not yet limiting the bandwidth. Otherwise, a discrepancy in bandwidth between on-resonance (purple) and off-resonance (red) operation should be observed. The kink at about 110 GHz is attributed to limitations in the electrical test set-up.

Two measurement set-ups were required to cover the broad electrical spectrum ranging from 20 GHz up to 115 GHz in steps of 2.5 GHz. The first set-up (20 GHz to 70 GHz) used a tunable radiofrequency (RF) signal generator (100 kHz to 70 GHz) to generate a sinusoidal RF signal. The electrical output power of the source was adjusted for each frequency such that a constant power of 0 dBm was applied to the ring resonator via ground–signal–ground probes. The second set-up was used to cover the frequency range from 70 GHz to 115 GHz with a step size of 2.5 GHz. We used the sixth harmonic of a Schottky diode to generate these frequencies. The electrical input power to the Schottky diode was kept constant. We compensated the variations in the diode's output power by measuring the power with a broadband power meter. The measured power values were used in a post-processing step to subtract the influence of the varying power. The power levels of the harmonic terms other than the sixth harmonic were negligible.

Note that the signal was applied via 100 μm × 100 μm contact pads. The resonator itself (approximately 1.1 fF) does not determine the total capacitance as the large contact pads interact with the substrate resulting in a capacitance of several tens of femtofarads[44]. The modulator behaves like an open circuit element owing to the small capacitance.

Data experiments. This section provides details about the experimental set-up shown in Fig. 4 of the main text. A tunable, external cavity laser (1,465 nm to 1,575 nm) was used as a light source with output powers ranging from 6 dBm up to 13 dBm. The laser light was fed to an attenuator to enable a flexible adjustment of the optical power. Light was coupled to and from the chip via grating couplers, which have an efficiency (C_{GC}) of −7 dB each at 1550 nm.

All non-return-to-zero (NRZ) electrical data streams were random binary sequences of length 2^{18} generated by a 72 gigasample s^{-1}, six-bit digital-to-analogue converter ($f_{3dB} \approx 20$ GHz). An RF amplifier ($f_{3dB} \approx 50$–55 GHz) then amplified the electrical signal to a certain peak-to-peak driving voltage ($V_{pp,50\Omega}$) measured over a 50 Ω resistor. Finally, the electrical signal was applied by ground-signal RF-probes ($f_{3dB} > 67$ GHz) to the device via 100 μm × 100 μm contact pads. Two 25-cm-long RF cables (1.85-mm connectors) connected the generator, amplifier and RF probe. The modulated light was detected coherently. We used a second laser for heterodyne detection. The 3 dB bandwidth of the coherent detector's photodiode is 70 GHz. An erbium-doped fibre amplifier amplified the optical signal after the modulator to power levels of about 8 dBm to enable best performance. The electrical oscilloscope (electrical bandwidth 63 GHz and sampling rate 160 GSa s^{-1}) sampled and digitized the waveform. Offline digital signal-processing was performed to account for timing errors[45], carrier frequency and phase offsets and the linear frequency response of the system by feedforward equalization with adaptive least mean square (LMS)-based filter updates. Extended Data Table 2 shows the results in terms of BER obtained from six different devices under various driving voltages (V_{peak}), data rates (R), and optical input powers at the grating coupler. Similar BERs are obtained for different devices under equal conditions. For instance, low BERs are obtained at 18 Gbit s^{-1} for small driving voltages of $V_{peak} \approx 3.3$ V and low input powers. At 72 Gbit s^{-1}, the BER drops to 10^{-3} under similar conditions. We increased the

driving voltages twofold and the input power fourfold and measured only a moderate reduction of the BER to 10^{-4}. We attribute this to the limited bandwidth of the electrical components. Device 34 is shown in the Fig. 4 of the main text.

Comparison with state-of-the-art electro-optic modulators. In the following, we compare the plasmonic ring resonators with prior art. Electro-optic-modulators are expected to help replace electrical communication links with their optical counterparts for short-range applications. Potential candidates must be compact (footprint of a few square micrometres) to enable a dense co-integration of electronics and photonics. Furthermore, electro-optic modulators must be optimized for high speed (> 100 GHz) and low insertion loss (< 3 dB) to enable multi-gigabit operation. Extended Data Fig. 6 shows the characteristics of state-of-the-art devices (photonic, crosses; plasmonic, circles) in terms of insertion loss and electro-optic-bandwidth[5,6,18,22,43,44,46–50]. The different abbreviations represent various electro-optic materials. Ideally, modulators should occupy the bluish area at the bottom right of the figure. However, large bandwidths correlate with large insertion loss. This is because the plasmonic waveguides also serve as highly conductive electrodes, enabling a low RC time constant, but ohmic losses in the metals prevent low insertion losses. Conversely, the resonant switching of photonic modulators enables low losses but these suffer from limited bandwidths. Combining plasmonics and the resonant switching concept breaks the trade-off and enables high-speed and low-loss modulators.

For the convenience of the reader, we provide a detailed summary of key metrics in Extended Data Table 3 (demonstrated plasmonic modulators) and in Extended Data Table 4 (photonic electro-optic modulators). The performance of the plasmonic ring resonator (last column) serves as reference in both tables. Extended Data Table 3 compares our modulator with four non-resonant modulators based on various nonlinear effects and material systems such as $\chi^{(2)}$ in OEO materials, free-carrier dispersion (FCD) in silicon, FCD in transparent conductive oxides (TCO) and $\chi^{(2)}$ in BTO.

• The resonant approach (last column) outperforms its direct non-resonant counterpart (first column).

• Plasmonic resonators requires moderate driving voltages, and only the TCO-based modulator is driven with $V_{peak} = 1$ V as required to enable integration with electronics. However, different data rates of 72 Gbit s^{-1} and 2.5 Gbit s^{-1} make a comparison between our modulator and the TCO modulator challenging.

• Finally, we note that applying the resonant plasmonic approach to other material systems would enable a further boost in performance for each individual technology. For instance, one can imagine TCO-based plasmonic resonators operating with millivolt driving voltages.

Extended Data Table 4 shows six photonic modulators based on FCD in silicon[18,46], the Franz–Keldysh effect (FKE) in SiGe (ref. [48]), $\chi^{(2)}$ in BTO[47], $\chi^{(2)}$ in lithium niobate (LiNbO$_3$)[50] and Pauli blocking in graphene[51].

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

33. Haffner, C. et al. Harnessing nonlinearities near material absorption resonances for reducing losses in plasmonic modulators. *Opt. Mater. Express* **7**, 2168–2181 (2017).
34. Jin, W. et al. Benzocyclobutene barrier layer for suppressing conductance in nonlinear optical devices during electric field poling. *Appl. Phys. Lett.* **104**, 243304 (2014).
35. Heni, W. et al. Nonlinearities of organic electro-optic materials in nanoscale slots and implications for the optimum modulator design. *Opt. Express* **25**, 2627–2653 (2017).
36. Kieninger, C. et al. In *Conf. on Lasers and Electro-Optics* STu3N.2 (Optical Society of America, 2017).
37. Maier, S. A. *Plasmonics: Fundamentals and Applications* (Springer, New York, 2007).
38. Messner, A. et al. In *Optical Fiber Communication Conf. Postdeadline Pap.* Th5C.7 (Optical Society of America, 2017).
39. Watanabe, T., Ayata, M., Koch, U., Fedoryshyn, Y. & Leuthold, J. Perpendicular grating coupler based on a blazed anti-back-reflection structure. *J. Lightwave Technol.* **35**, 4663–4669 (2017).
40. Kretz, J., Dreeskornfeld, L., Ilicali, G., Lutz, T. & Weber, W. Comparative study of calixarene and HSQ resist systems for the fabrication of sub-20nm MOSFET device demonstrators. *Microelectron. Eng.* **78/79**, 479–483 (2005).
41. Elder, D. L. et al. Effect of rigid bridge-protection units, quadrupolar interactions, and blending in organic electro-optic chromophores. *Chem. Mater.* **29**, 6457–6471 (2017).
42. Gheorma, I. L. & Osgood, R. M. Fundamental limitations of optical resonator based high-speed EO modulators. *IEEE Photon. Technol. Lett.* **14**, 795–797 (2002).
43. Hoessbacher, C. et al. Plasmonic modulator with 170 GHz bandwidth demonstrated at 100 GBd NRZ. *Opt. Express* **25**, 1762–1768 (2017).
44. Melikyan, A. et al. High-speed plasmonic phase modulators. *Nat. Photonics* **8**, 229–233 (2014).
45. Josten, A. et al. Modified Godard timing recovery for non integer oversampling receivers. *Appl. Sci.* **7**, 655 (2017).

46. Alloatti, L., Cheian, D. & Ram, R. J. High-speed modulator with interleaved junctions in zero-change CMOS photonics. *Appl. Phys. Lett.* **108**, 131101 (2016).
47. Girouard, P. et al. chi2 modulator with 40-GHz modulation utilizing BaTiO$_3$ photonic crystal waveguides. *IEEE J. Quantum Electron.* **53**, 1–10 (2017).
48. Srinivasan, A. et al. In *Optical Fiber Communication Conf.* Tu3D.7 (Optical Society of America, 2016).
49. Haffner, C. et al. Plasmonic organic hybrid modulators: scaling highest speed photonics to the microscale. *Proc. IEEE* **104**, 2362–2379 (2016).
50. Wang, C., Zhang, M., Stern, B., Lipson, M. & Loncar, M. Nanophotonic lithium niobate electro-optic modulators. Preprint at https://arxiv.org/abs/1701.06470 (2017).
51. Phare, C. T., Daniel Lee, Y.-H., Cardenas, J. & Lipson, M. Graphene electro-optic modulator with 30 GHz bandwidth. *Nat. Photon.* **9**, 511–514 (2015).
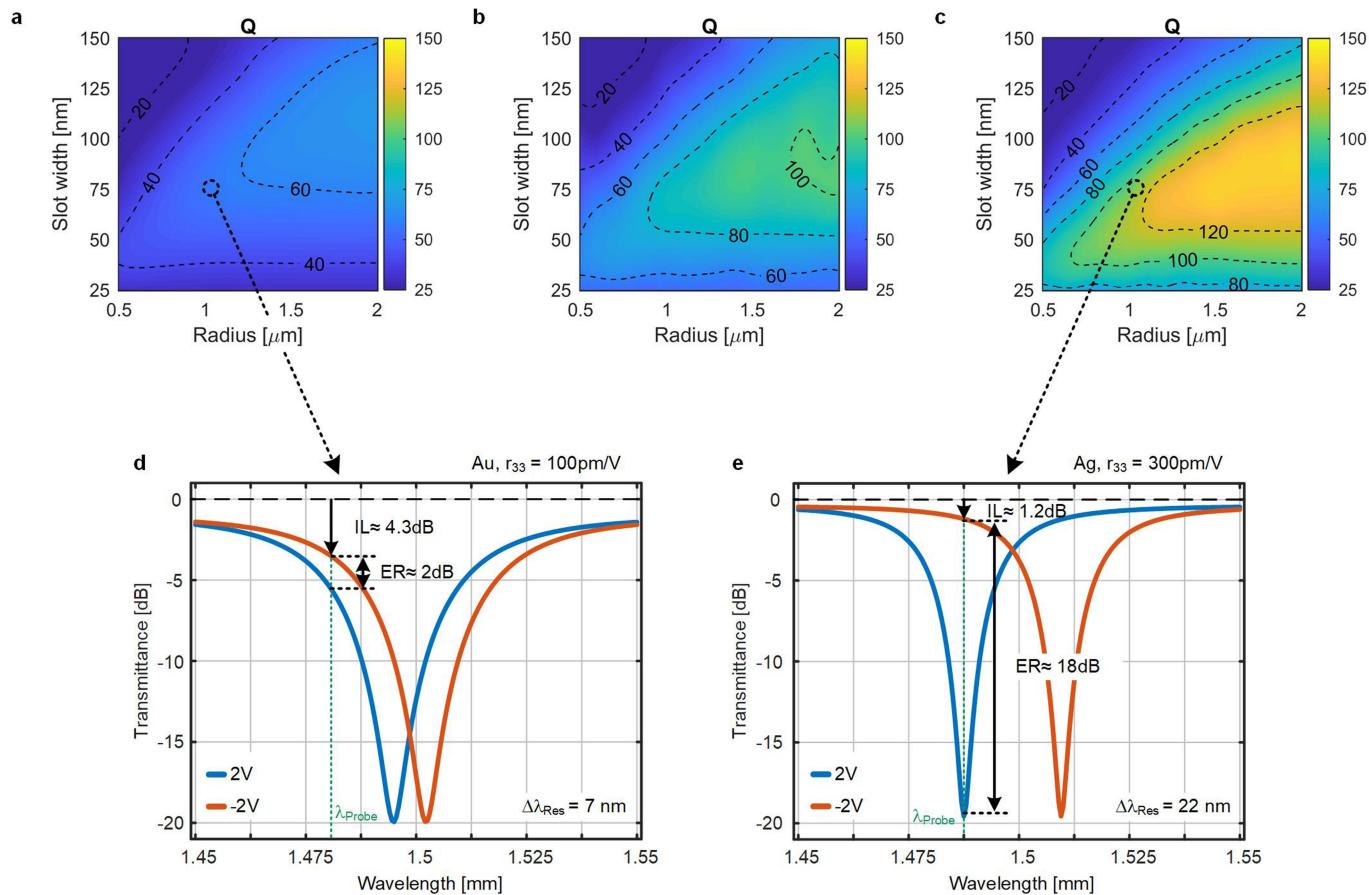
Extended Data Fig. 1 | Resonator performance for various plasmonic materials. a, Gold, which is interesting for research because of its chemical stability. **b**, Copper is of interest as it is a CMOS-compatible material. **c**, Silver features the best plasmonic properties and could be of interest for high-performance applications. **d**, **e**, Switching capability of (**d**) Au and (**e**) Ag ring resonator for a 2 V bias. The latter uses the newest OEO material, which has a three times larger electro-optic coefficient, r_{33}. The performance improvement enables a considerable reduction in terms of the driving voltage. The number at the bottom right indicates the shift in the resonance wavelength.
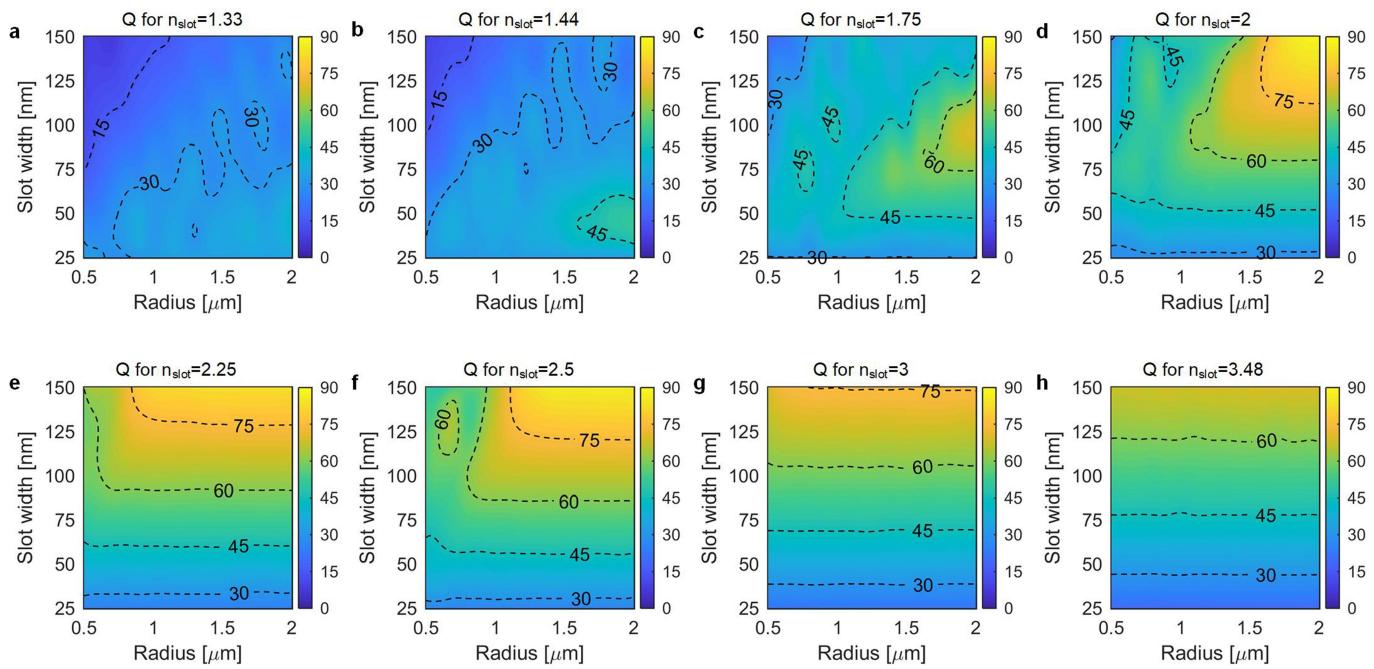
Extended Data Fig. 2 | Q-factors of various materials filling the slot.
a–h, The materials differ in their refractive index, and one can observe that low-n materials are limited by bending loss (diagonal lines) whereas high-n materials are limited by propagation loss (parallel lines). These simulations were performed with 150 nm height of the outer and inner electrode to account for limitations in fabrication processes different from ours.

SU8200 5.0kV x15.0k SE(U)   3.00μm
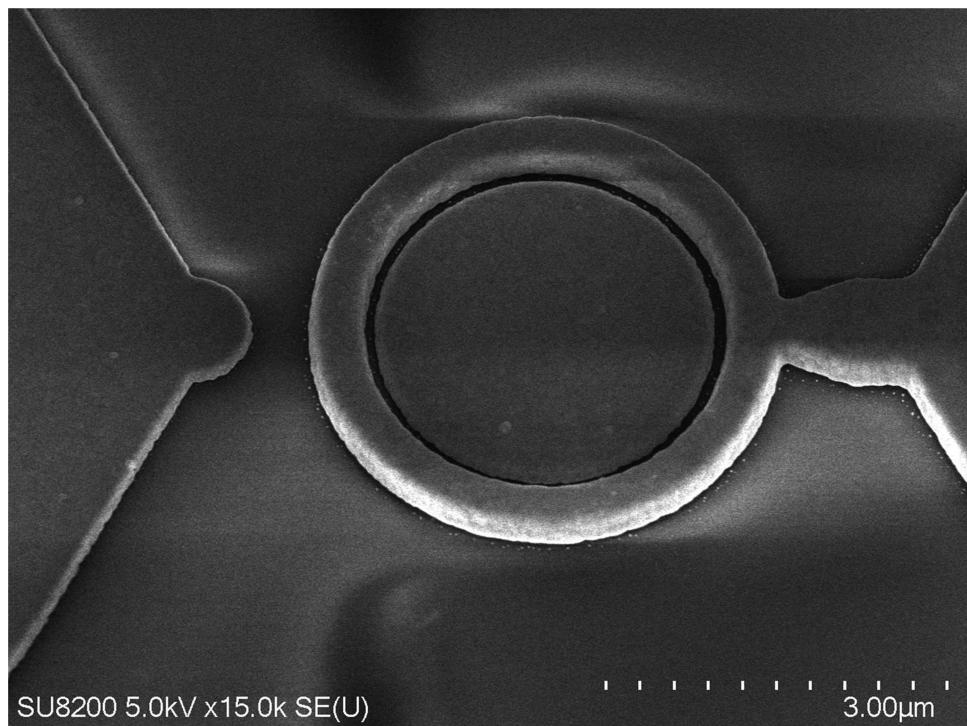
Extended Data Fig. 3 | Tilted SEM image of a processed ring resonator. The different height of the outer and inner electrodes reduces the bending losses.

a

b

c
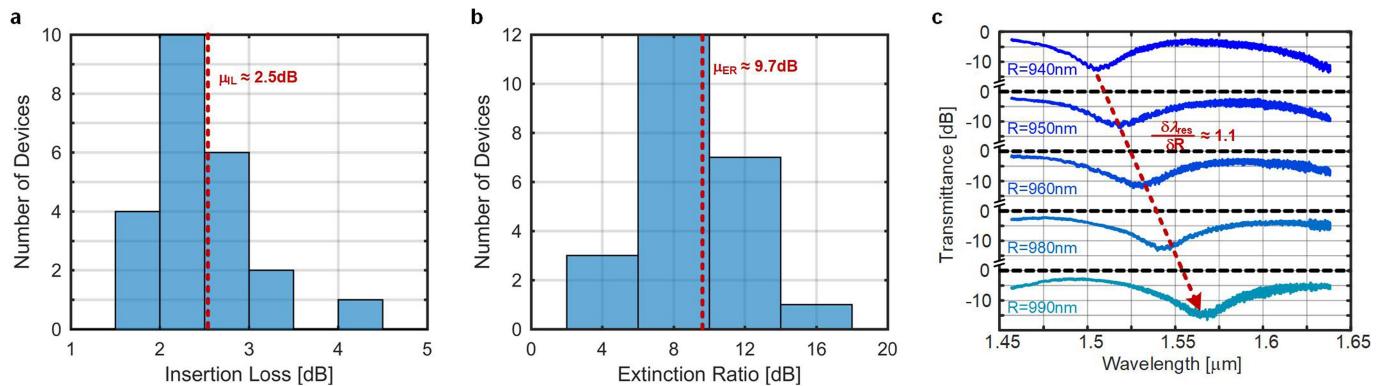
Extended Data Fig. 4 | Reproducibility of plasmonic ring resonators.
a, Insertion loss and **b**, extinction ratio histograms. Data are obtained from passive measurements of 23 devices with a designed slot width of 80 nm and radii ranging from 900 nm to 1,100 nm. **c**, Dependence of the resonance wavelength on ring radius.
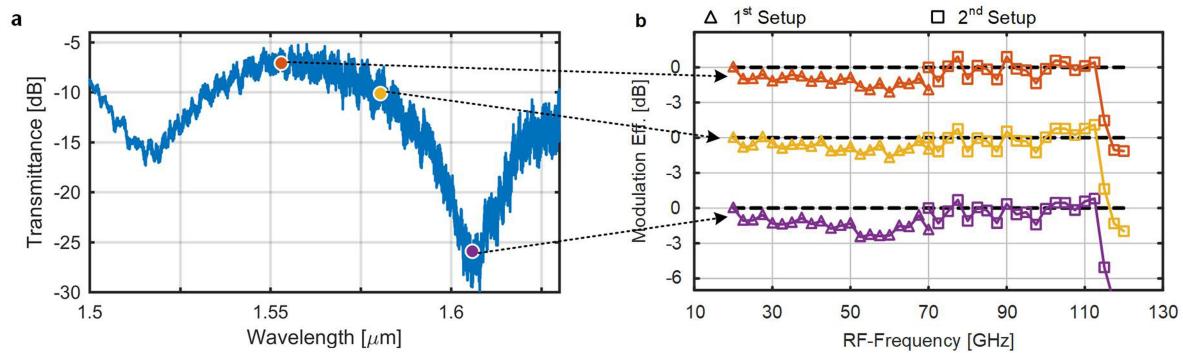
Extended Data Fig. 5 | Transmission spectrum and the measured bandwidth at the off-resonance, 3 dB and on-resonance operating point. a, Transmission spectrum; **b**, measured bandwidth. No bandwidth limitation can be observed up to 110 GHz. The drop at 115 GHz frequencies is due to a limited measurement set-up. Recent studies show that the modulation efficiency at lower radiofrequency is not limited[44].
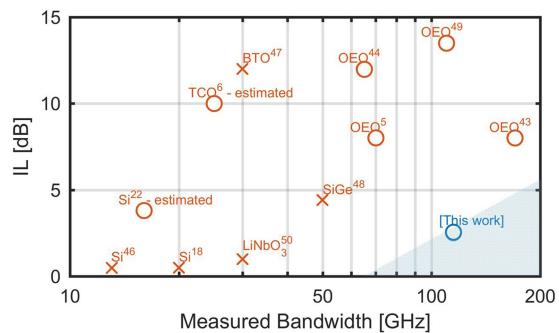
Extended Data Fig. 6 | Technology overview in terms of insertion loss and bandwidth of electro-optic modulators. Ideal candidates should feature low insertion loss with high electro-optic bandwidths.

Extended Data Table 1 | Measured d.c. sensitivity

Radius [nm]/ Slot width [nm], (Device Id)	*IL* [dB]	*ER* [dB]	$\Delta\lambda_{res}$/Bias [nm/V]
1200/80, (22)	~3	7	3
1260/80, (28)	~3	8	N.A.
1220/100, (36)	~2	8	1.8
1250/100, (39)	~2	6	2.1
1280/100, (42)	~2.5	7.5	2.33
1080/80, (79)	~2.75	10	2.75

The values of the radius and the slot width are the design parameters.

Extended Data Table 2 | Overview of the results obtained from data experiments with various devices

Radius [nm]/ Slot width [nm], Device Id	V_{peak} [V]	Data rate (R) [Gbit/s]	Bit-error-ratio (*BER*)	$P_{opitcal}$ [dBm]
1230/80, (25)	~3.3	18	~9×10^{-6}	0
1250/80, (27)	~3.3	18	<1×10^{-6}	0
1200/100, (34)	~3.3	18	<1×10^{-6}	0
"	~3.3	36	~2×10^{-5}	4
"	~3.3	72	~1×10^{-3}	4
960/80, (88)	~5	72	~2×10^{-3}	6
"	~7.8	72	~8×10^{-4}	6
940/80, (86)	~7.8	72	~4×10^{-4}	12
930/80, (85)	~7.8	72	~1×10^{-3}	12

Extended Data Table 3 | Comparison with state-of-the-art plasmonic electro-optic modulators

Reference	43	22	6	38	this work
Modulation Mech.	$\chi^{(2)}$	FCD	FCD	$\chi^{(2)}$	$\chi^{(2)}$
	OEO	Si	TCO	BTO	OEO
IL [dB]	~8	~3.8	~10	25	~2.5
Fiber-2-fiber loss[dB]	34	N.A.	N.A.	N.A.	16
ER [dB]	N.A.	6-10	6	>15	10
L_{Active} [μm]	20	1	4	10	6
Bandwidth [GHz]	>170	16*	>2.5	N.A.	>110
Data Rate [Gbit/s]	100	<<1	2.5	72	72
Driving Voltage [V_{peak}]	4	6	1	2.8	3.3
Energy Consumption [fJ/Bit]	48*	~100	N.A.	N.A.	12

*Estimated values

Extended Data Table 4 | Comparison with state-of-the-art photonic electro-optic modulators

Reference	18	46	47	50	51	48	this work
Modulation Mech.	FCD Si	FCD Si	$\chi^{(2)}$BTO	$\chi^{(2)}$LiNbO$_3$	Pauli Blocking Graphene	FKE Si Ge	$\chi^{(2)}$OEO
IL [dB]	~ 0.5	~0.5	>12	1	N.A.	4.4	~ 2.5
ER [dB]	12	10	N.A.	~7*	30	4.2	10
L$_{active}$ [μm]	15	30	37	~600*	~250	40	6
Bandwidth [GHz]	~17	13	30*	30	30	>50	>110
Data Rate [Gbit/s]	44	12	N.A.	40	22	50	72
Driving Voltage [V$_{peak}$]	0.5	2.46	16	2.88	3.75	1	3.3
Energy Consumption [fJ/Bit]	<1	N.A.	N.A.	240	800	13.8	12
Q	6000	15000	N.A.	8000	<1200	N.A	50
Temp. Stability [$\Delta\lambda_{res}$ K^{-1} FWHM^{-1}]	100%	N.A.	N.A.	N.A.	N.A.	N.A	0.4%
Sensitivity [$\Delta\lambda_{res}$ V^{-1} FWHM^{-1}]	100%	11%	N.A.	N.A.	N.A.	N.A	5%

LETTER

Earthquake–induced transformation of the lower crust

Bjørn Jamtveit[1]*, Yehuda Ben-Zion[2], François Renard[1,3] & Håkon Austrheim[1]

The structural and metamorphic evolution of the lower crust has direct effects on the lithospheric response to plate tectonic processes involved in orogeny, including subsidence of sedimentary basins, stability of deep mountain roots and extension of high-topography regions. Recent research shows that before orogeny most of the lower crust is dry, impermeable and mechanically strong[1]. During an orogenic event, the evolution of the lower crust is controlled by infiltration of fluids along localized shear or fracture zones. In the Bergen Arcs of Western Norway, shear zones initiate as faults generated by lower-crustal earthquakes. Seismic slip in the dry lower crust requires stresses at a level that can only be sustained over short timescales or local weakening mechanisms. However, normal earthquake activity in the seismogenic zone produces stress pulses that drive aftershocks in the lower crust[2]. Here we show that the volume of lower crust affected by such aftershocks is substantial and that fluid-driven associated metamorphic and structural transformations of the lower crust follow these earthquakes. This provides a 'top-down' effect on crustal geodynamics and connects processes operating at very different timescales.

The structural and metamorphic evolutions of the lower crust are key elements in the dynamics of the lithosphere. Frequent observations of fluid-induced metamorphism associated with ductile deformation along shear zones on scales ranging from millimetres to kilometres inspired early models of the lithosphere such as the 'jelly-sandwich' model[3]. In this model, the lower crust is assumed to be wet and mechanically weak, and plate tectonic stress is transmitted through the brittle upper crust and a strong upper mantle. This model was challenged[4] with the argument that a strong lower crust is essential for the survival of thick mountain roots and high mountains. The lower crust is dominated by granulite facies rocks of mafic to intermediate composition[5] and such rocks will be nominally dry at normal steady-state geothermal gradients for a wide range of crustal heat flow and heat production conditions[1]. Hence, the rheology of the lower crust before an orogeny will in most cases be controlled by the properties of dry mineral assemblages dominated by plagioclase, pyroxene, garnet and olivine, with plagioclase being the most abundant phase. This is consistent with the estimated viscosity of the lower crust ($\geq 10^{24}$ Pa s) required to generate the crustal support needed for intraplate seismicity such as the Bhuj earthquake[6] in western India in 2001 of moment magnitude $M_w = 7.6$. Recent modelling[7] furthermore suggests that the Indian lower crust remains strong beneath the entire southern half of the Tibetan plateau.

Observations of structural and metamorphic transformation of initially dry lower crust during orogenic events indicate an early stage involving seismic failures[8–10]. Metamorphism and shear zone development then follow in the wake of lower-crustal earthquakes. These observations raise an enigma that has so far been unresolved, because frictional failure of dry rocks at the confining pressures of the lower crust (≥ 1 GPa) requires differential stress levels[11] exceeding 2 GPa. Although dry plagioclase-dominated rocks deforming by dislocation creep can in theory develop extremely high differential stress at

lower-crustal temperatures and high strain rates, the stress level that can be sustained over orogenic timescales for reasonable strain rates in coherent crustal volumes will be far below what is required for brittle faulting (< 1 GPa; see Methods). Deep crustal earthquakes occurring under constant loading in intact rocks thus seem to require a local weakening mechanism.

During subduction of the Indian plate under south Tibet, earthquakes occur at 60–100 km depth (Fig. 1), but are confined to regions very close to the Mohorovičić discontinuity (Moho) at temperatures[7] below 600 °C. Interestingly, the pressure and temperature conditions in the region where these earthquakes nucleate overlap the conditions at which serpentine breaks down to produce hydrous fluids in mantle rocks. Fluid production near the Moho can both reduce effective pressures and weaken the crust and mantle by mineral transformation processes and thus it is a plausible explanation for the observed seismic activity.

In the absence of such local weakening mechanisms, seismic deformation in the lower crust may be driven by transient 'stress pulses'[6,10,12]. Here we propose that earthquakes in the brittle upper crust provide a natural mechanism for sustained generation of stress pulses and associated seismic failures in the lower crust. During the occurrence of large earthquakes the strain rates around and below the rupture area increase by many orders of magnitude. A representative strain accumulation of 15 mm per year across a width comparable to a geodetic locking depth of 15 km corresponds to an interseismic strain rate of 3×10^{-14} s^{-1}. In contrast, seismic slip velocities of 1–10 m s^{-1} across a rupture localization width of 1–10 mm lead to seismic strain rates of 10^3–10^4 s^{-1}. Such large co-seismic jumps can explain a transient increase in seismic rupture within the lower crust.

Observational evidence for very high, short-lived stresses in the lower crust come from the occurrence of fossil earthquakes. A recent study of the Woodroffe Thrust located within the Musgrave Block in central Australia[13] documents the formation of large volumes of pseudotachylytes in completely dry lower-crustal granulites. Stress levels exceeding 0.5 GPa have also been inferred from lower-crustal earthquakes, leading to pseudotachylyte formation in gabbros and ultramafic rocks in the Alpine subduction complex of Corsica, France[14].

Simulated deformation on faults using various versions of rate- and state-dependent friction models show that large earthquake slip penetrates into the nominally stable deeper regions of the crust[15,16]. Simulations of aftershocks in a viscoelastic damage model consisting of a brittle upper crust over a lower crust with power-law viscosity constrained by laboratory experiments indicate that the hypocentres of early aftershocks are deeper than the regular seismogenic zone[2]. Depending on model parameters and thermal gradients, the maximum depth of the early aftershocks can approach twice that of the usual seismicity. Details of these results depend on the constitutive laws and parameters used, but lower-crustal aftershocks are generic outcomes of the high strain rates generated by large earthquakes at the bottom of the seismogenic zone.

[1]Physics of Geological Processes (PGP), The Njord Centre, Department of Geosciences, University of Oslo, Oslo, Norway. [2]Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA. [3]Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, Grenoble, France. *e-mail: bjorn.jamtveit@geo.uio.no
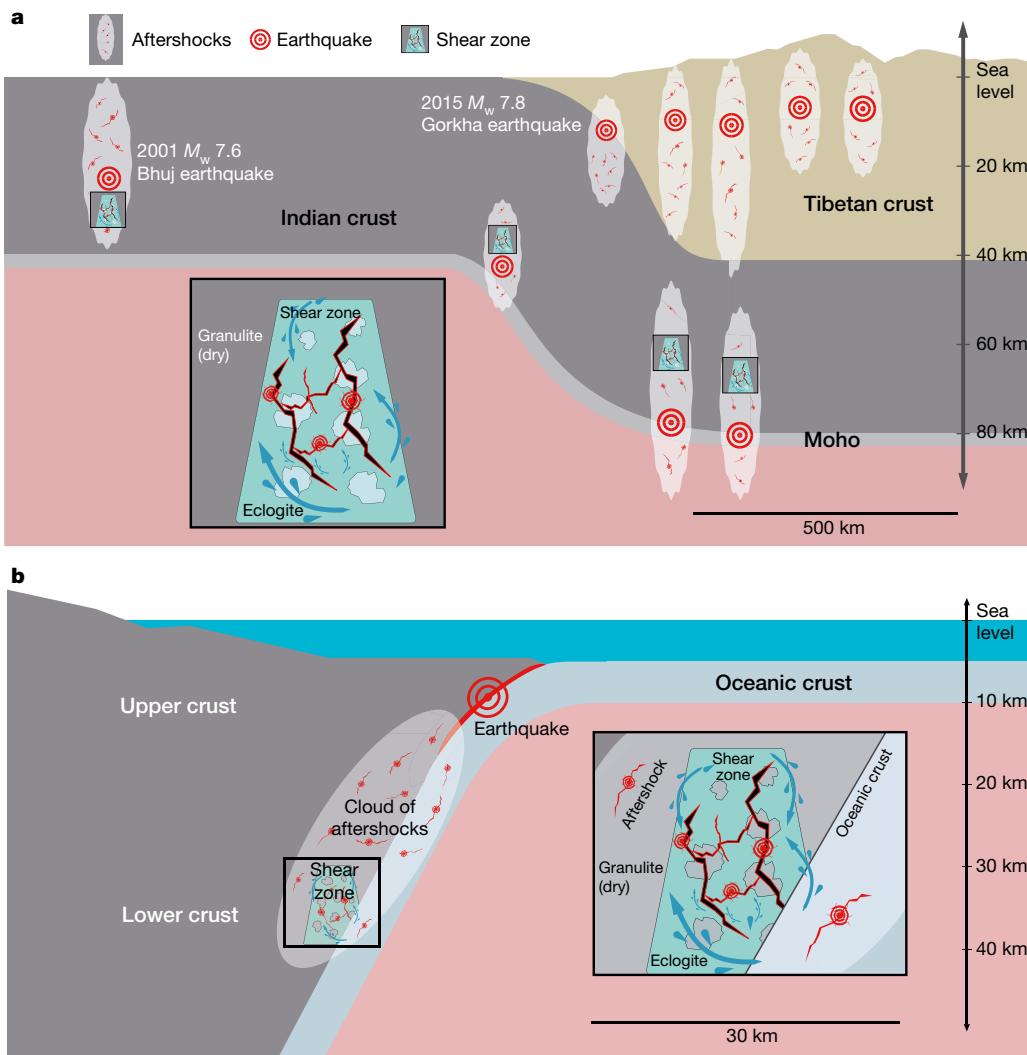
Fig. 1 | Earthquakes and aftershocks in the lower crust. a, Schematic representation of earthquakes and aftershocks for the India–Tibet continent–continent collision. Inset, enlarged shear zone. **b**, A generic subduction plate boundary geometry. Each major earthquake generates a cloud of aftershocks, some of which are in the lower crust (many aftershocks in the upper crust are not shown in these simple diagrams). These aftershocks create pathways for fluids (blue arrows in insets), allowing partial hydration and metamorphism of the strong and dry granulites into wet and weaker eclogites and amphibolites. This process also facilitates the development of shear zones in the continental lower crust. For the subduction geometry, fluids could originate from the slab below or from the upper crust above. For the continental collision below the Himalayas, fluids introduced to the subducted Indian plate could originate from the dehydration of serpentine rocks below. The 2001 Bhuj[6] and 2015 Gorkha[28] earthquakes both have aftershock 'clouds' propagating down to the lower crust, potentially allowing downward migration of fluids from the upper crust.

Below we use basic scaling relations to demonstrate that observed worldwide earthquake activity in the regular seismogenic zone of subduction zones and seismically active continental regions is expected to produce considerable fracture area and rupture zone volume in the lower crust. This, in turn, generates transient pathways for fluids from the wet upper crust above, or the slab below, to the dry lower crust (Fig. 1). Fluids have a key role in the long-term evolution of the lower crust[1]. As an example, we describe earthquake-triggered eclogite-facies metamorphism and shear zone development of lower-crustal granulites from the Bergen Arcs in Western Norway (Fig. 2a). The observations highlight the close association between earthquakes, fluid migration and transformation processes in the lower crust.

The Bergen Arcs represent a series of thrust sheets where granulite facies remnants of Proterozoic lower crust recrystallized to an anhydrous mineralogy 930 million years ago[17]. During the Caledonian continent collision between Laurentia and Baltica between 420 and 440 million years ago[17], fluid-induced metamorphic transformations formed eclogites and amphibolites in shear zones, breccias and along fractures. The estimated eclogitization conditions[18] are about 670–690 °C and 2.1–2.2 GPa. Pseudotachylytes, fine-grained or glassy fault rocks believed to reflect earthquake-related frictional melting are abundant on faults where granulites facies rocks experienced Caledonian retrograde metamorphism (Fig. 2a). Such faults show single rupture displacements reaching 1.7 m (Fig. 2b), corresponding to an earthquake exceeding magnitude 7. Single pseudotachylyte veins range in thickness from millimetres to a few centimetres (Fig. 2c), and also occur as a thin 'matrix' between rotated blocks of brecciated granulite that sometimes cover areas exceeding 100 m². Microstructures developing in the fault wall rocks display intense fragmentation with no, or very minor, shear strain[19], followed by healing processes through grain growth and formation of eclogite facies minerals, including hydrous phases such as amphibole, mica and clinozoisite. Infiltration of hydrous fluids is thus directly associated with seismic slip.

A substantial rheological weakening associated with formation of the fine-grained and hydrous eclogite often leads to development of ductile shear zones in areas initially deformed by brittle failure. Relict pseudotachylytes can occasionally be observed 'floating' in the shear

a, Map (modified from ref. [33], Elsevier) of the northwestern part of Holsnøy island in the Bergen Arcs, Western Norway, showing the location of pseudotachylytes (red stars) recording numerous fossil earthquakes near the transition between 940-million-year-old dry lower-crustal granulites and hydrated 430-million-year-old eclogites. **b**, Offset of a pyroxene-rich seam by a single lower-crustal earthquake. The slip surface is decorated with a melt layer (pseudotachylyte) indicative of seismic slip. An offset of 1.7 m corresponds to a fossil earthquake with moment magnitude $M_w \geq 7$ based on the scaling relations provided in ref. [29]. **c**, Higher-resolution image of the centimetres-thick melt layer between the two red dashed lines. Red arrows indicate the sense of slip. Photographs in **b** and **c** taken by B.J.

Fig. 2 | Fossil earthquakes in the Bergen Arcs.

Fig. 3 | Transformation of the lower crust. Offset of a pseudotachylyte by a shear zone where dry granulite rocks are transformed into wet eclogites. The earthquake occurred before the eclogitization and probably created the pathways for fluids that triggered rock transformation. We note the remains of the pseudotachylyte inside the eclogite shear zone. Photograph taken by F.R.

zones, providing unambiguous evidence for ductile deformation pre-dated by brittle failure of granulite facies rocks (Fig. 3). In the following, we explore the feasibility that lower-crustal earthquakes, such as those described, are aftershocks triggered by stress pulses generated by mainshocks in the normal seismogenic regime of a plate boundary.

Basic seismological scaling relations provide an order-of-magnitude estimate of the lower-crustal rock volume affected by aftershocks. We show that this is substantial, with conservative parameter values and ignoring probable contributions from penetration of large mainshocks into the lower crust as well as ductile/thermal instabilities[15,16].

Lower-crustal earthquakes are not expected to occur repeatedly in the same location because rock melting and subsequent solidification is a strengthening process[20]. This is consistent with observations of distributed 'fields' of pseudotachylytes (Fig. 2a), each associated with

a single event. The total volume of rock damage V_{rd} produced by crustal earthquakes in the magnitude range $M_1 < M < M_2$ is then given by

$$V_{rd} = \int_{M_1}^{M_2} A(M) \, t_{rd}(M) \, n(M) \, dM \tag{1}$$

Here $A(M)$ is the rupture area of an earthquake with magnitude M, $t_{rd}(M)$ is the damage zone thickness around the rupture area and $n(M)$ is the event density given by the Gutenberg–Richter relationship

$$\log_{10} n(M) = a - bM \tag{2}$$

where a and b are empirical constants. The scaling relations for $A(M)$ and $t_{rd}(M)$ are found using basic theoretical relations from fracture mechanics[21,22] and empirical relation between the magnitude and potency of earthquakes[23,24] (see Methods). An explicit relationship for the total volume of damage produced by earthquakes in the considered magnitude range can be expressed as:

$$V_{rd}(M_1 < M < M_2) = \beta[e^{\alpha M_2} - e^{\alpha M_1}] \tag{3}$$

where α and β are positive parameters that account for the combined scaling relations[21–24] of A, t_{rd} and n with M. By using observationally constrained parameters, the volume of rock damaged by crustal earthquakes is estimated to be $V_{rd} = 1.2 \times 10^{-5}$ km^3 per year per square kilometre of the Earth's surface in the seismically active region (see Methods). For the lower-magnitude limit, we use $M_1 = 0$ with slip distance of the order of the grain size of granulites. For an upper limit relevant to the lower crust, we take $M_2 = 8.3$ since the largest subduction zones events can have $M \geq 9.5$, and the largest aftershock magnitude is typically about 1.2 units below that of the mainshock[25,26]. The parameters of the Gutenberg–Richter relationship are taken from recent analysis of global earthquakes with depths[27] of less than 70 km. Analysis of earthquake clusters indicates[26,27] that about 75% of all events with $M > 0$ are aftershocks. We therefore use 75% of the observed intensity of events[27] to estimate the average annual production of damaged rock volume by aftershocks with $0 < M < 8.3$. On the basis of previously conducted simulations[2], we assume conservatively that 1% of the aftershock population is in the lower crust. The estimated annual production of rupture zone volume in the lower crust is then 1.2×10^{-7} km^3 yr^{-1} per square kilometre of the Earth's surface.

As a concrete example, Western Norway was an active subduction zone during the Caledonian orogeny for more than two million years. From the estimates above, the total seismically damaged volume in the lower crust of Western Norway is estimated to be $0.24\,km^3$ per square kilometre of the Earth's surface. For a 20-km-thick lower crust, this implies a rupture zone volume exceeding 1.2% of the total lower-crustal volume. The Bergen Arcs example demonstrates that in the presence of fluids, lower-crustal earthquakes initiate metamorphism of rock volumes typically 1–2 orders of magnitude larger than that of the rupture zone (that is, a 0.1–1-m-thick eclogite forming around the 1-cm-thick rupture zone). Hence, the overall process can alter a large fraction of the lower crust. The sensitivity of the results to input parameters is discussed in the Methods and shows that the 1.2% estimate of lower-crust volume damaged by earthquakes is based on conservative values of input parameters and is likely to be higher.

Our results indicate that aftershocks triggered by major earthquakes in the regular seismogenic zone have the potential to initiate pervasive transformation of the lower crust on a timescale of a million years. Direct recording of a transient deepening of early aftershocks requires a dense observational network around large mainshock ruptures. Although this situation is not often met, such lower-crust aftershocks are sometimes observed. Recent examples of deep aftershocks include the 2001 $M_w = 7.6$ Bhuj intraplate earthquake in India where aftershocks occurred down to Moho depths[6], and the 2015 $M_w = 7.8$ Gorkha earthquake in Nepal where the hypocentre occurred near the Main Himalayan Thrust at a depth[28] of 10–15 km and aftershocks penetrated the Indian crust to a depth exceeding 30 km (Fig. 1a).

Earthquakes in dry lower crust driven by stress pulses generated in the seismogenic zone have a number of important consequences, many of which are illustrated by the Bergen Arcs example. The most important is arguably the associated increase in permeability, which may connect the dry lower crust to an external fluid reservoir. In the Bergen Arcs, pseudotachylyte formation is always associated with an influx of hydrous fluids[29]. Fluid-driven metamorphic reactions are fast owing to the metastable state of the granulite facies rocks[1], leading to a profound reduction in rock strength and the development of shear zones and ductile deformation at lower stress levels[30,31]. The positive feedbacks between fluid introduction, weakening and shear zone development eventually produce a complete transformation of large volumes of lower crust from a dry and strong lithology to a wet and weak lithology. Thus even in situations where aftershocks directly affect only limited volumes of lower crust, they may start a series of fluid-induced transformation processes which can affect far bigger volumes.

The 'top-down' control on lower-crustal evolution presented here challenges the traditional 'bottom-up' view in which deep shear zones are assumed to control the spatial distribution of faults above the brittle–ductile transition. The generation of deep crustal shear zones as a response to weakening induced by pre-existing faults triggered by stress pulses generated by shallower earthquakes may also explain observed fluids with meteoric and other upper crustal signatures, such as the presence of hydrocarbons, in shear zones formed below the brittle–ductile transition[32].

Online content

1. Jamtveit, B., Austrheim, H. & Putnis, A. Disequilibrium metamorphism of stressed lithosphere. *Earth Sci. Rev.* **154**, 1–13 (2016).
2. Ben-Zion, Y. & Lyakhovsky, V. Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology. *Geophys. J. Int.* **165**, 197–210 (2006).
3. Chen, W. P. & Molnar, P. Focal depths of intracontinental and intraplate earthquakes and their implication for the thermal and mechanical properties of the lithosphere. *J. Geophys. Res.* **88**, 4183–4214 (1983).
4. Jackson, J. Strength of the continental lithosphere: time to abandon the jelly sandwich? *GSA Today* **12**, 4–9 (2002).
5. Rudnick, R. L. & Fountain, D. M. Nature and composition of the continental lower crust. *Rev. Geophys.* **33**, 267–309 (1995).
6. Copley, A., Avouac, J.-P., Hollingsworth, J. & Leprince, S. The 2001 M_W 7.6 Bhuj earthquake, low fault friction and the upper crustal support of plate driving forces in India. *J. Geophys. Res.* **116**, B08405 (2011).
7. Craig, T. J., Copley, A. & Jackson, J. Thermal and tectonic consequences of India underthrusting Tibet. *Earth Planet. Sci. Lett.* **353–354**, 231–239 (2012).
8. Austrheim, H. & Boundy, T. M. Pseudotachylytes generated during seismic faulting and eclogitization of the deep crust. *Science* **265**, 82–83 (1994).
9. John, T. & Schenk, V. Interrelations between intermediate-depth earthquakes and fluid flow within subducting oceanic plates: constraints from eclogite facies pseudotachylytes. *Geology* **34**, 557–560 (2006).
10. Moecher, D. P. & Steltenpohl, M. G. Direct calculation of rupture depth for an exhumed paleoseismogenic fault from mylonitic pseudotachylyte. *Geology* **37**, 999–1002 (2009).
11. Kohlstedt, D. L., Evans, B. & Mackwell, S. J. Strength of the lithosphere—constraints imposed by laboratory experiments. *J. Geophys. Res.* **100**, 17587–17602 (1995).
12. Ellis, S. & Stöckhert, B. Elevated stresses and creep rates beneath the brittle-ductile transition caused by seismic faulting in the upper crust. *J. Geophys. Res.* **109**, B05407 (2004).
13. Hawemann, F., Mancktelow, N. S., Wex, S., Camacho, A. & Pennacchioni, G. Pseudotachylyte as field evidence for lower crustal earthquakes during the intracontinental Petermann Orogeny (Musgrave Block, Central Australia). *Solid Earth Discuss.* (submitted); preprint at https://doi.org/10.5194/se-2017-123 (2018).
14. Andersen, T. B., Mair, K., Austrheim, H., Podladchikov, Y. Y. & Vrijmoed, J. C. Stress release in exhumed intermediate and deep earthquakes determined from ultramafic pseudotachylite. *Geology* **36**, 995–998 (2008).
15. Hillers, G., Ben-Zion, Y. & Mai, P. M. Seismicity on a fault controlled by rate- and state dependent friction with spatial variations of the critical slip distance. *J. Geophys. Res.* **111**, B01403 (2006).
16. Jiang, J. & Lapusta, N. Connecting depth limits of interseismic locking, microseismicity, and large earthquakes in models of long-term fault slip. *J. Geophys. Res.* **122**, 6491–6523 (2017).
17. Bingen, B., Davis, W. J. & Austrheim, H. Zircon U-Pb geochronology in the Bergen arc eclogites and their Proterozoic protoliths, and implications for the pre-Scandian evolution of the Caledonides in western Norway. *Geol. Soc. Am. Bull.* **113**, 640–649 (2001).
18. Bhowany, K. et al. Phase equilibria modelling constraints on P-T conditions during fluid catalysed conversion of granulite to eclogite in the Bergen Arcs, Norway. *J. Metamorph. Geol.* **36**, 315–342 (2018).
19. Austrheim, H. et al. Microstructural records of seismic slip. *Sci. Adv.* **3**, e1602067 (2017).
20. Mitchell, T. M., Toy, V., Di Toro, G., Renner, J. & Sibson, R. H. Fault welding by pseudotachylyte formation. *Geology* **44**, 1059–1062 (2016).
21. Ben-Zion, Y. Collective behavior of earthquakes and faults: continuum-discrete transitions, progressive evolutionary changes and different dynamic regimes. *Rev. Geophys.* **46**, RG4006 (2008).
22. Ben-Zion, Y. & Ampuero, J.-P. Seismic radiation from regions sustaining material damage. *Geophys. J. Int.* **178**, 1351–1356 (2009).
23. Ben-Zion, Y. & Zhu, L. Potency-magnitude scaling relations for southern California earthquakes with $1.0 < M_L < 7.0$. *Geophys. J. Int.* **148**, F1–F5 (2002).
24. Edwards, B., Allmann, B., Fah, D. & Clinton, J. Automatic computation of moment magnitudes for small earthquakes and the scaling of local to moment magnitude. *Geophys. J. Int.* **183**, 407–420 (2010).
25. Båth, M. Lateral inhomogeneities in the upper mantle. *Tectonophysics* **2**, 483–514 (1965).
26. Zaliapin, I. & Ben-Zion, Y. Y. Earthquake clusters in southern California. I: Identification and stability. *J. Geophys. Res.* **118**, 2847–2864 (2013).
27. Zaliapin, I. & Ben-Zion, Y. A global classification and characterization of earthquake clusters. *Geophys. J. Int.* **207**, 608–634 (2016).
28. McNamara, D. E. et al. Source modeling of the 2015 M_W 7.8 Nepal (Gorkha) earthquake sequence: implications for geodynamics and earthquake hazards. *Tectonophysics* **714–715**, 21–30 (2017).
29. Wells, D. L. & Coppersmith, K. J. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull. Seismol. Soc. Am.* **84**, 974–1002 (1994).
30. Putnis, A., Jamtveit, B. & Austrheim, H. Metamorphic processes and seismicity: the Bergen Arcs as a natural laboratory. *J. Petrol.* **58**, 1871–1898 (2017).
31. Yardley, B. W. D. The role of water in the evolution of the continental crust. *J. Geol. Soc. Lond.* **166**, 585–600 (2009).
32. Munz, I. A., Yardley, B. W. D., Banks, D. & Wayne, D. Deep penetration of sedimentary fluids in basement rocks from Southern Norway—evidence from hydrocarbon and brine inclusions in quartz veins. *Geochim. Cosmochim. Acta* **59**, 239–254 (1995).
33. Austrheim, H. Fluid and deformation induced metamorphic processes around Moho beneath continent collision zones: Examples from the exposed root zone of the Caledonian mountain belt, W-Norway. *Tectonophysics* **609**, 620–635 (2013).

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0045-y.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to B.J.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Controls on lower-crustal stress levels. We used the flow law creep parameters for synthetic plagioclase aggregates for variable water contents from ref. [34]. Extended Data Fig. 1 shows the relations between temperature, strain rate and differential stress for water-poor plagioclase aggregates based on these experimental data. At 660–680 °C, the estimated temperature during seismic faulting in the Bergen Arcs, the maximum differential stress developing for a strain rate of $10^{-14}\,\mathrm{s}^{-1}$ would be about 0.3 GPa, and clearly insufficient to cause brittle failure. Such strain rates would probably only apply within zones where strain was already localized. In a coherent crustal volume, the internal strain rate would be far less. Strain rates are estimated[7] at around $10^{-16}\,\mathrm{s}^{-1}$ for the subducting Indian plate beneath eastern Tibet. Even if local strain rate increases should be able to push stress levels beyond what would be expected in coherent crustal slabs, or if the crustal temperature was lower than for the Bergen case so that the effective viscosity was higher, stresses much higher than 1 GPa would be unrealistic owing to the onset of Peierls creep[35], or even cataclastic flow[36].

Volume of damage produced by earthquakes. The step-by-step derivation of equation (3) and the calculation of the volume of rock damaged by earthquakes in the crust, V_{rd}, are detailed here. The total volume of rock damage produced by crustal earthquakes in the magnitude range $M_1 < M < M_2$ is given by equation (1). Each term in the integral of equation (1) is described below.

The density of events $n(M)$ in equation (1) is provided by the Gutenberg–Richter relationship, equation (2), where a and b are empirical constants. Assuming that each earthquake is approximately a circular rupture with radius r and surface area $A = \pi r^2$ that sustains a uniform strain drop $\Delta\varepsilon$, the seismic potency P_0 (moment divided by rigidity) is given by[21]

$$P_0 = (16/7)\,\Delta\varepsilon\, r^3 \tag{4}$$

The seismic potency and magnitude of earthquakes spanning a relatively small range (≤ 4) of magnitudes are related empirically by a relation of the type[23,24]

$$\log_{10} P_0 = dM + e \tag{5}$$

with constants d and e.

Combining equations (4) and (5), the radius of an earthquake with magnitude M in units of kilometres is:

$$r(M) = \left[\frac{7 \times 10^{-5}}{16\Delta\varepsilon}\right]^{1/3} \times 10^{(dM+e)/3} \tag{6}$$

where the 10^{-5} factor accounts for the unit conversion of P_0 from $\mathrm{km}^2\,\mathrm{cm}$ to km^3. Using equation (6) for r and assuming again a circular crack such that $A(M) = \pi r^2$, we obtain

$$A(M) = \pi\left[\frac{7 \times 10^{-5}}{16\Delta\varepsilon}\right]^{2/3} \times 10^{2(dM+e)/3} \tag{7}$$

The thickness of the damage zone t_{rd} is expected from fracture mechanics to scale linearly with the rupture radius[22].

$$t_{\mathrm{rd}}(M) = \gamma r(M) \tag{8}$$

where the constant γ is proportional to the dynamic stress intensity factor and the ratio of stress drop over strength drop. For standard rupture velocity of 0.9 times the Rayleigh wave speed and a stress drop that is 10% of the strength drop, $\gamma \approx 10^{-2}$–10^{-3} (situations with relatively high initial stress leading to higher stress drop give higher γ values). Here we conservatively use $\gamma = 0.002$.

Combining equations (2), (7) and (8) leads to

$$A(M)\,t_{\mathrm{rd}}(M)\,n(M) = \pi\gamma\left[\frac{7 \times 10^{-5}}{16\Delta\varepsilon}\right] \times 10^{(d-b)M+e+a} \tag{9}$$

Integrating equation (1) using equation (9) gives an explicit relationship for the total volume of damage produced by earthquakes—that is, equation (3), where

$\alpha = (d-b)\ln(10)$ and $\beta = \frac{\pi}{\alpha}\gamma 10^{a+e}\left(\frac{7 \times 10^{-5}}{16\Delta\varepsilon}\right)$ are positive parameters and V_{rd} is in units of cubic kilometres per year per square kilometre of the Earth's surface in a seismically active region.

Observed b-values are typically around 1, while a-values vary considerably with space and time. We focus on deformation in subduction zones and seismically active continental regions and use average a and b values based on analysis of global earthquakes with depth shallower than 70 km during 1975–2015 in the Northern California Earthquake Data Center (NCEDC) catalogue[27]. Using figure 1 of ref. [27] and a b-value of 1, a representative intensity of events with $M > 0$ (10^a) in active subduction zones is 6 per year per square kilometre, corresponding to $a = \log_{10}(6) = 0.78$. This value is conservative because the NCEDC catalogue does not include numerous small events buried in the coda of larger events and noise. The results of ref. [27] indicate that about 50% of $M > 4$ events in subduction zones and continental transform regions are aftershocks. The fraction of events that are aftershocks increases as the event magnitude decreases[26,27], so we assume that about 75% of all events with $M > 0$ are aftershocks. The event intensity 10^a used below to estimate damaged rock volume by aftershocks is reduced accordingly by a factor of 0.75 from 6 to 4.5.

Analysis of earthquakes in southern California recorded by the regional network and borehole sensors indicates[23] that $d = 1$ and $e = -4.7$ for $M < 3.5$ while $d = 1.34$ and $e = -5.22$ for $M > 3.5$. Similar constants characterize earthquakes in other locations[24]. Inserting these constants d and e for magnitude ranges $0 < M < 3.5$ and $3.5 < M < 8.3$ into equation (9) and integrating, a more explicit expression of the volume of damaged rocks (in cubic kilometres per year per square kilometre of the Earth's surface at the seismically active region) is

$$V_{\mathrm{rd}} = \int_{M_1}^{M_2} A(M)\,t_{\mathrm{rd}}(M)\,n(M)\,\mathrm{d}M = \pi\gamma\left[\frac{7 \times 10^{-5}}{16\Delta\varepsilon}\right] \times 10^a \int_{M_1}^{M_2} 10^{(d-b)M+e}\mathrm{d}M \tag{10}$$

and

$$V_{\mathrm{rd}} = \pi\gamma\left[\frac{7 \times 10^{-5}}{16\Delta\varepsilon}\right]10^a\left\{\int_0^{3.5} 10^{-4.7}\mathrm{d}M + \int_{3.5}^{8.3} 10^{0.34M} \times 10^{-5.22}\mathrm{d}M\right\} \tag{11}$$

Evaluating equations (10) and (11) gives $V_{\mathrm{rd}} = 1.2 \times 10^{-5}\,\mathrm{km}^3$ per year per square kilometre of the Earth's surface in a seismically active region.

The fraction of lower-crust volume affected by earthquake ruptures (1.2%) is obtained using $10^a = 4.5$, $\Delta\varepsilon = 5 \times 10^{-5}$ and the values of other constants mentioned in the text. The result is sensitive to the input parameters, but it is based on realistic values of earthquake intensities in active subduction zones (4.5 annual aftershocks with $M > 0$ per square kilometre), assumed lower and upper limits of aftershock magnitudes (0 and 8.3), average strain drop ($\Delta\varepsilon = 5 \times 10^{-5}$), ratio of damage zone thickness to rupture radius $\gamma = 0.002$, and the fraction of aftershocks with hypocentres in the lower crust (1%). Reducing the lower magnitude limit will increase the rupture surface area but not change the estimated damage zone volume much; decreasing the upper magnitude limit from 8.3 to 7.8 will decrease the estimated volume of damaged rock by 32%. Changing the assumed average strain drop $\Delta\varepsilon = 5 \times 10^{-5}$ to average strain drops of 5×10^{-4} and 5×10^{-6} will modify the estimated damage volume by factors of 0.21 and 4.64, respectively. Changing γ or the deformation timescale by a given factor (for example, 2) will modify the damage volume by the same factor. Thus, reasonable variations of these parameters will have only minor effects on the estimated lower-crust volume affected by earthquake ruptures.

Data availability. All of the data used are contained within the paper.

34. Rybacki, E. & Dresen, G. Deformation mechanism maps for feldspar rocks. *Tectonophysics* **382**, 173–187 (2004).
35. Azuma, S., Katayama, I. & Nakakuki, T. Rheological decoupling at the Moho and implications to Venusian tectonics. *Sci. Rep.* **4**, 4403 (2014).
36. Tullis, J. & Yund, R. The brittle–ductile transition in feldspar aggregates. An experimental study. *Int. Geophys.* **51**, 89–117 (1992).
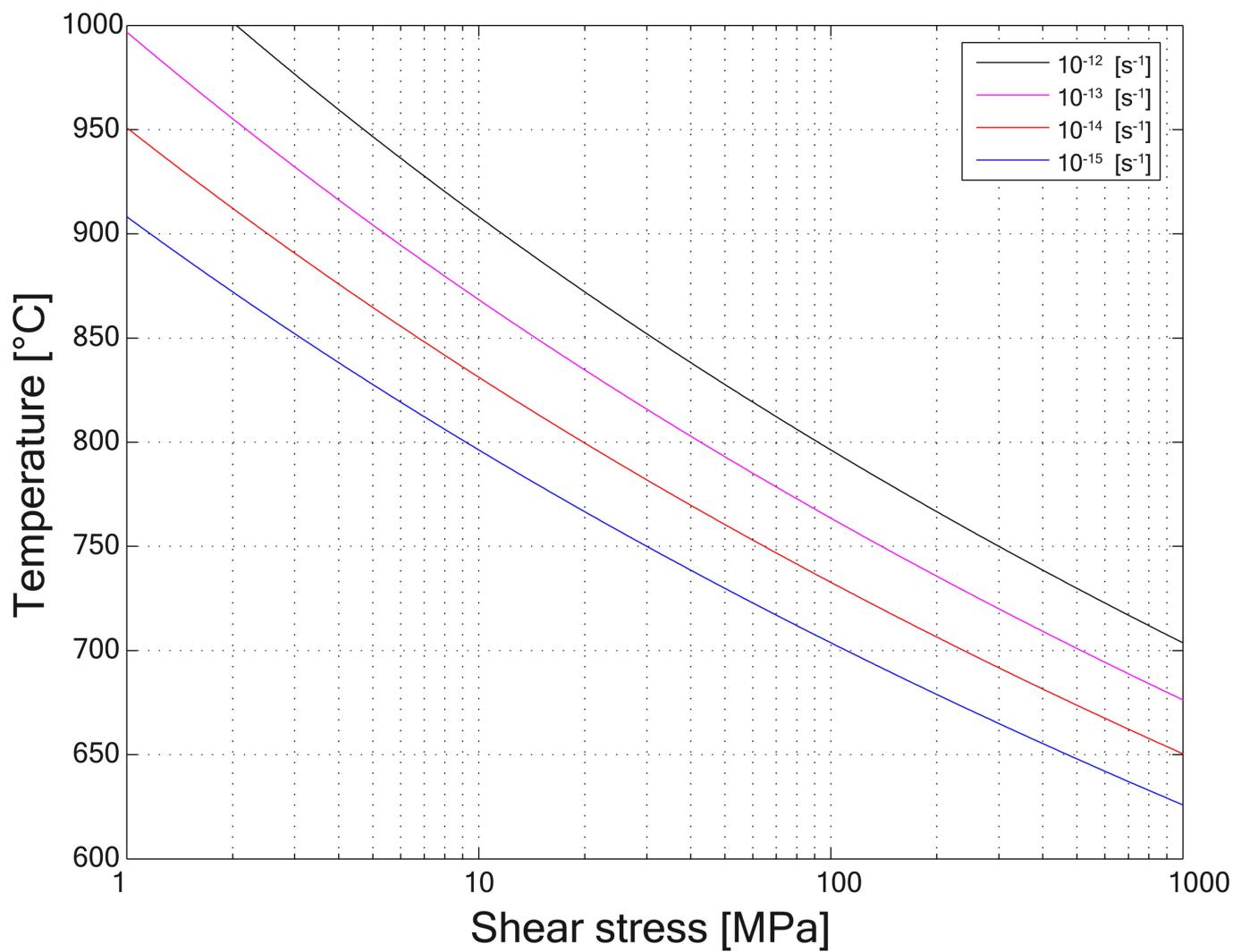
Extended Data Fig. 1 | Rheology of dry anorthite. Shear stress versus temperature diagram contoured with respect to strain rate.

Global warming transforms coral reef assemblages

Terry P. Hughes[1]*, James T. Kerry[1], Andrew H. Baird[1], Sean R. Connolly[1,2], Andreas Dietzel[1], C. Mark Eakin[3], Scott F. Heron[3,4,5], Andrew S. Hoey[1], Mia O. Hoogenboom[1,2], Gang Liu[3,4], Michael J. McWilliam[1], Rachel J. Pears[6], Morgan S. Pratchett[1], William J. Skirving[3,4], Jessica S. Stella[6] & Gergely Torda[1,7]

Global warming is rapidly emerging as a universal threat to ecological integrity and function, highlighting the urgent need for a better understanding of the impact of heat exposure on the resilience of ecosystems and the people who depend on them[1]. Here we show that in the aftermath of the record-breaking marine heatwave on the Great Barrier Reef in 2016[2], corals began to die immediately on reefs where the accumulated heat exposure exceeded a critical threshold of degree heating weeks, which was 3–4 °C-weeks. After eight months, an exposure of 6 °C-weeks or more drove an unprecedented, regional-scale shift in the composition of coral assemblages, reflecting markedly divergent responses to heat stress by different taxa. Fast-growing staghorn and tabular corals suffered a catastrophic die-off, transforming the three-dimensionality and ecological functioning of 29% of the 3,863 reefs comprising the world's largest coral reef system. Our study bridges the gap between the theory and practice of assessing the risk of ecosystem collapse, under the emerging framework for the International Union for Conservation of Nature (IUCN) Red List of Ecosystems[3], by rigorously defining both the initial and collapsed states, identifying the major driver of change, and establishing quantitative collapse thresholds. The increasing prevalence of post-bleaching mass mortality of corals represents a radical shift in the disturbance regimes of tropical reefs, both adding to and far exceeding the influence of recurrent cyclones and other local pulse events, presenting a fundamental challenge to the long-term future of these iconic ecosystems.

Extreme weather events due to anthropogenic global warming are rapidly emerging as major contemporary threats to almost all ecosystems[1]. On coral reefs, severe heatwaves trigger episodes of mass bleaching[4–7], which occur when the relationship between corals and their photosynthetic symbionts (zooxanthellae, *Symbiodinium* spp.) breaks down, turning the coral pale. Bleached corals are physiologically damaged and nutritionally compromised, and they can die if the bleaching is severe and the recovery time of their symbionts is prolonged[8,9]. However, the relationship between heat exposure, bleaching and the initial and longer term mortality of different taxa is not well understood or quantified. Although the concept of winners versus losers has been widely applied to describe inter-specific differences in the degree of bleaching[10–14], predicting the definitive losers, namely those corals that fail to regain their colour and ultimately die following heat stress, is key to understanding how climate change affects biodiversity, species composition and ecosystem function. To date, no study has, to our knowledge, examined the quantitative relationship between a broad range of heat exposures and the response of coral assemblages. Establishing the shape of this response curve is essential for identifying the critical levels of heat exposure that trigger bleaching and mass mortality, and for predicting the amount of heat exposure that could drive a transformation in species composition and the widespread collapse of ecological functions. Here, we examine geographical patterns of heat exposure and the resultant mortality of coral assemblages along the 2,300 km

length of the Great Barrier Reef, following the record-breaking marine heatwave of 2016[2]. We show that taxonomic patterns of bleaching did not predict the identity of the corals that ultimately died, that many corals succumbed immediately from heat stress, and that others died more slowly following the depletion of their zooxanthellae. The die-off of corals drove a radical shift in the composition and functional traits of coral assemblages on hundreds of individual reefs, transforming large swaths of the Great Barrier Reef from mature and diverse assemblages to a highly altered, degraded system.

The 2016 bleaching event triggered an unprecedented loss of corals on the northern third of the Great Barrier Reef, and to a lesser extent, the central third, with almost no heat-stress mortality occurring further south (Fig. 1a and Extended Data Figs. 1–3). The geographical footprint and intensity of the coral die-off (Fig. 1a) closely matched the observed north–south pattern in accumulated heat (Fig. 1b), measured as satellite-derived degree heating weeks (DHW in °C-weeks), a commonly used measurement that incorporates both the duration and intensity of heat stress[15,16]. The 5-km-resolution DHW values (Fig. 1b) were significantly correlated with independently estimated losses of corals (Fig. 1a; $r^2 = 0.50$, $P < 0.001$, $n = 1,156$ reefs). In the northern, 700-km-long section of the Great Barrier Reef (from 9.5–14.5 °S), in which the heat exposure was the most extreme, 50.3% of the coral cover on reef crests was lost within eight months (Fig. 1b). More broadly, throughout the entire Great Barrier Reef, including the southern third, in which the heat exposure was minimal (Fig. 1b), the cover of corals declined by 30.0% between March and November 2016. In comparison, the massive loss of corals from the 2016 marine heatwave was an order of magnitude greater and more widespread than the patchier, localized damage that typically occurs on reef sites within the track of a severe tropical cyclone[17].

At the scale of individual reefs, the severity of coral mortality was also highly correlated with the amount of bleaching, and with the level of heat exposure (Fig. 2). Initially, at the peak of temperature extremes in March 2016, many millions of corals died quickly in the northern third of the Great Barrier Reef over a period of only 2–3 weeks (Fig. 2a). These widespread losses were not due to the attrition of corals that slowly starved because they failed to regain their symbionts[9]. Rather, temperature-sensitive species of corals began to die almost immediately in locations that were exposed to heat stress of more than 3–4 °C-weeks (Figs. 1b, 2a). The amount of initial mortality increased steadily with increasing heat exposure ($r^2 = 0.50$, $P < 0.001$, $n = 63$ reefs); on reefs which were exposed to less than 4 °C-weeks, fewer than 5% of the corals died, whereas an initial median loss of 15.6% of corals was recorded on reefs with 4–8 °C-weeks exposure, and a median loss of 27.0% of corals at locations that experienced 8 °C-weeks or more (Fig. 2a). Across the entire Great Barrier Reef, 34.8% of individual reefs experienced at least 4 °C-weeks, and 20.7% of reefs were exposed to 8 °C-weeks or more of accumulated heat stress in 2016 (Fig. 1b). The amount of initial mortality at the peak of summer varied strikingly among different groups of corals (Extended Data Fig. 4a).

[1]Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland, Australia. [2]College of Science and Engineering, James Cook University, Townsville, Queensland, Australia. [3]Coral Reef Watch, US National Oceanic and Atmospheric Administration, College Park, MD, USA. [4]Global Science & Technology, Inc., Greenbelt, MD, USA. [5]Marine Geophysical Laboratory, Physics Department, James Cook University, Townsville, Queensland, Australia. [6]Great Barrier Reef Marine Park Authority, Townsville, Queensland, Australia. [7]Australian Institute of Marine Science, Townsville, Queensland, Australia. *e-mail: terry.hughes@jcu.edu.au
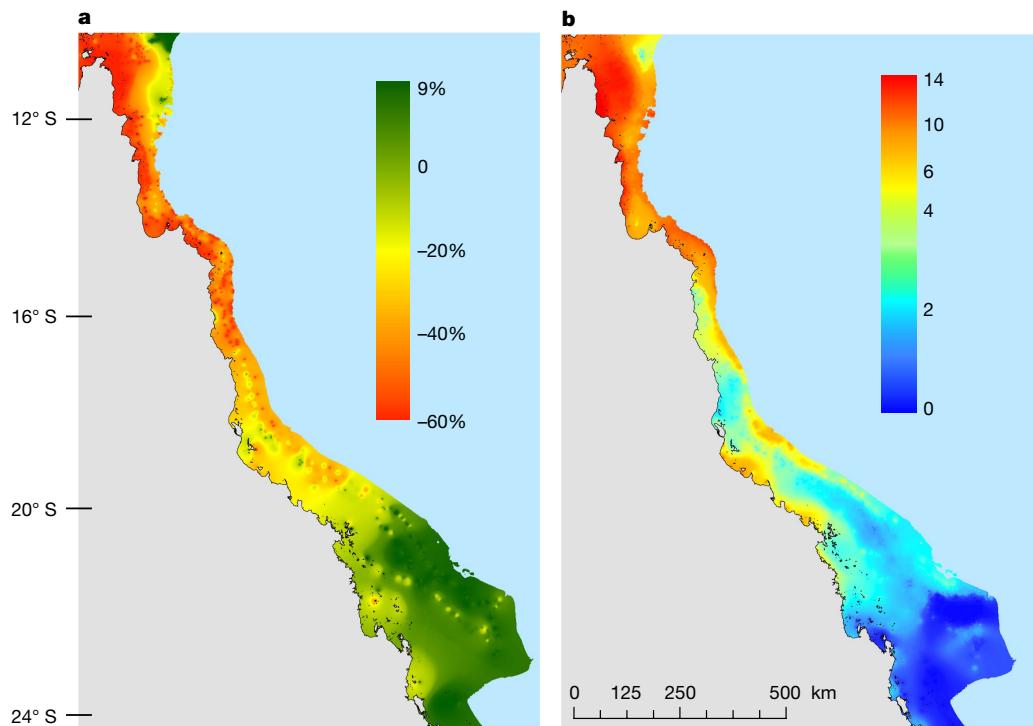
Fig. 1 | Large-scale spatial patterns in change in coral cover and in heat exposure on the Great Barrier Reef, Australia. a, Change in coral cover between March and November 2016. **b**, Heat exposure, measured in DHW (in °C-weeks) in the summer of 2016. Map template is provided by Geoscience Australia (© Commonwealth of Australia (Geoscience Australia) 2018).

During the ensuing Austral winter, the bleached corals in the northern and central Great Barrier Reef either slowly regained their colour and survived or they continued to die at unprecedented levels. Less than 1% of surviving colonies remained bleached after eight months. The severity of the longer term loss of corals, measured in situ as the decline in coral cover between March and November, was accurately predicted by the percentage of corals that were initially bleached (Fig. 2b; $r^2 = 0.51$, $P < 0.001$, $n = 63$ reefs). Specifically, reefs that experienced less than 25% bleaching in March typically had almost no loss of cover after eight months (Fig. 2b). By contrast, above this threshold, the loss of coral cover increased progressively, indicating that fewer of the bleached corals survived. Furthermore, the longer term loss of coral cover also intensified with increasing levels of heat exposure (DHW)

experienced by each reef ($r^2 = 0.44$, $P < 0.001$, $n = 63$ reefs; Fig. 2c). Consequently, we recorded almost no loss of coral cover for reefs exposed to 0–3 °C-weeks, compared to a 40% decline at 4 °C-weeks, 66% for 8 °C-weeks, and extreme declines of > 80% for exposures of 9 °C-weeks or more. The nonlinear responses to heat exposure varied significantly among coral taxa (Extended Data Figs. 5, 6), illustrating a spectrum of survivorship among winners versus losers, driving a radical shift in species composition.

Post-bleaching mortality has disproportionately transformed the assemblage structure and functional diversity of corals on reefs that experienced high levels of bleaching (affecting more than 60% of colonies), as illustrated by a non-metric multi-dimensional scaling (nMDS) analysis (Fig. 3). The abundances of all categories of corals
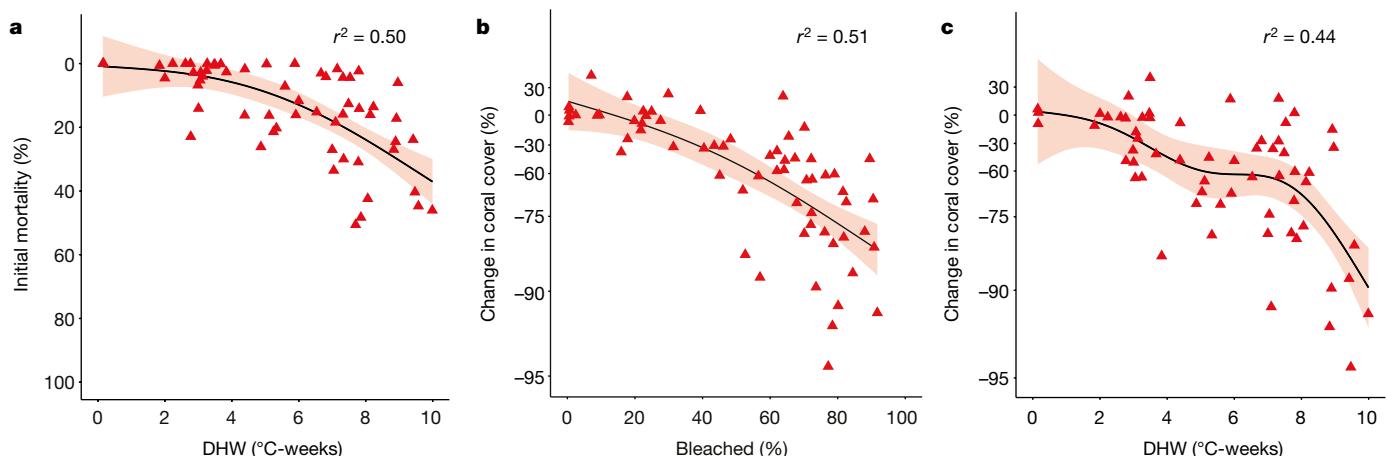
Fig. 2 | The initial and longer term response of coral assemblages to heat exposure. Regression curves were fitted using generalized additive models, with 95% confidence limits (ribbons). Data points represent individual reefs. **a**, Initial coral mortality measured at the peak of bleaching ($n = 63$ reefs), versus the heat exposure each reef experienced (satellite-based DHW (in °C-weeks)). **b**, Longer term change in coral cover ($\log_{10}$) between March and November 2016 on 63 individual reefs, versus the initial amount of bleaching recorded underwater. **c**, Longer term change in coral cover ($\log_{10}$) between March and November 2016, versus heat exposure (DHW) on the same individual reefs.

Assemblage structure
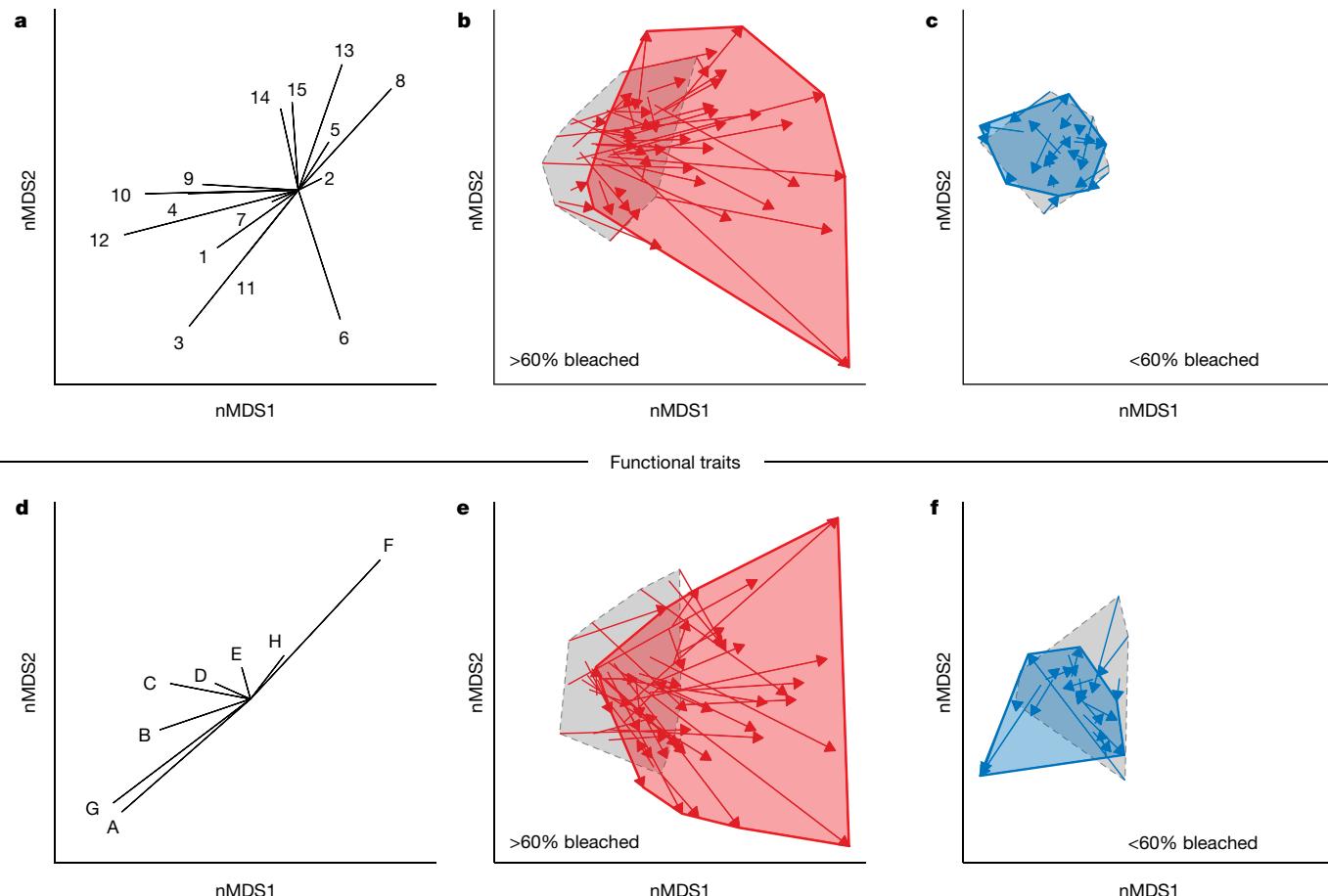
Functional traits

Fig. 3 | Changes in assemblage structure and functional traits of corals following mass bleaching. a–c, nMDS analyses of shifts in coral assemblages between March and November 2016. **a**, Fifteen nMDS vectors indicate the responses of individual taxa: 1, other *Acropora*; 2, favids; 3, *Isopora*; 4, *Montipora*; 5, Mussidae; 6, other *Pocillopora*; 7, *Pocillopora damicornis*; 8, Poritidae; 9, *Seriatopora hystrix*; 10, staghorn coral (*Acropora* spp.); 11, *Stylophora pistillata*; 12, tabular coral (*Acropora* spp.); 13, soft corals; 14, other scleractinia; 15, other sessile fauna (see Methods). **b**, The grey polygon bounds the ordination space occupied by coral assemblages on each reef in March. Red arrows connect the before–after pairs of data points for each location to show changes in composition on severely bleached reefs (> 60% of colonies bleached, *n* = 43 reefs) after eight months (in November), bounded by the red polygon. **c**, Blue arrows

connect the before–after pairs of data points for each location on reefs (*n* = 20) that were moderately (< 60% bleached), bounded by the grey (March) and blue polygons (November). **d–f**, nMDS analyses of shifts in assemblage trait composition between March and November 2016 at the same locations. **d**, The eight vectors indicate the absolute contribution of traits to coral assemblages: A, surface area to volume ratio; B, growth rate; C, colony size; D, skeletal density; E, colony height; F, corallite width; G, interstitial space size; H, reproductive mode (see Methods and Extended Data Table 1). **e**, The shift in abundance-weighted trait space coordinates for coral assemblages over eight months for reefs with > 60% bleaching. **f**, The shift in abundance-weighted trait space coordinates for coral assemblages on reefs with < 60% bleaching.

decreased to varying degrees on these heavily bleached reefs, shown by the orientation of the nMDS vectors (Fig. 3a) and the directional shift in the before–after assemblages (Fig. 3b). Tabular and staghorn *Acropora*, *Seriatopora hystrix* and *Stylophora pistillata*—fast-growing, three-dimensional species that dominate many shallow Indo-Pacific reefs—all declined by > 75% (Extended Data Fig. 4b). In contrast to the radical shifts on heavily bleached reefs, assemblages changed very little between March and November on reefs that experienced moderate (30–60%) or minor (0–30%) bleaching (Fig. 3c).

The response of coral assemblages on reefs exposed to a broad range of heat stress, ranging from 0 to 10 °C-weeks, was strikingly nonlinear (Fig. 4). The changes in assemblage structure after eight months (measured as the Euclidean distance between before and after compositions on each reef; Fig. 3b, c) were small on reefs that were exposed to less than 6 °C-weeks, whereas reefs subjected to more than 6 °C-weeks lost over 50% of their corals (Fig. 2c) and shifted markedly in composition (Fig. 4). Satellite-derived DHW data indicate that 28.6% of the 3,863 reefs comprising the Great Barrier Reef experienced thermal exposures of more than 6 °C-weeks during the 2016 bleaching event, and 20.7%

(800 reefs) were exposed to more than 8 °C-weeks (Fig. 1). Individual reefs with this severity of heat exposure have undergone an unprecedented ecological collapse, extending southwards from Papua New Guinea for up to 1,000 km (Fig. 1). Reefs that were exposed to less than 6 °C-weeks were located predominantly in the southern half of the Great Barrier Reef, and in a narrow northern patch at the outer edge of the continental shelf where temperature anomalies in 2016 above the local long-term summer maximum were small (Fig. 1b).

The abrupt, regional-scale shift in coral assemblages has also radically reduced the abundance and diversity of species traits that facilitate key ecological functions (Fig. 3d, e and Extended Data Tables 1, 2). A before–after analysis of the multi-dimensional trait space of coral assemblages, weighted by the absolute abundance of taxa contributing to each trait, reveals a transformation in the functional-trait composition of assemblages on heavily bleached reefs (affecting over 60% of colonies) in the eight-month period after March 2016 (Fig. 3e). In most cases, reefs shifted away from the dominance of fast-growing, branching and tabular species that are important providers of three-dimensional habitat, to a depauperate assemblage dominated by taxa
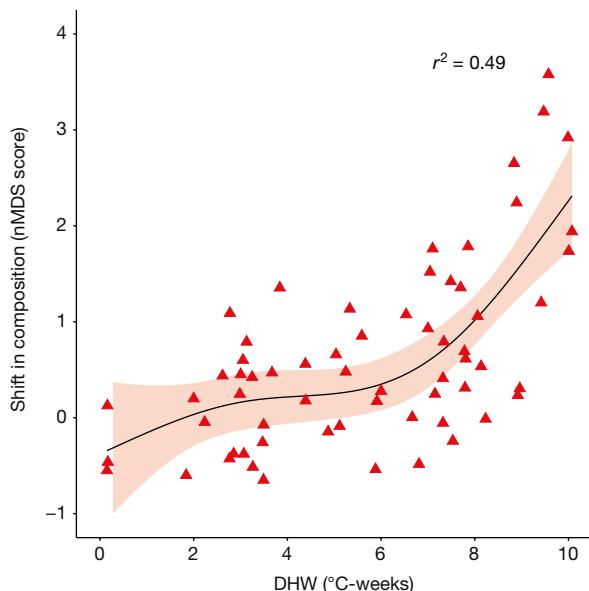
Fig. 4 | Change in coral assemblages in response to heat exposure.
The regression curve is fitted using a generalized additive model, with
95% confidence limits. Each data point represents the shift in composition
($n = 63$ reefs), based on the Euclidean distance in a non-metric multi-
dimensional scaling analysis of assemblages on individual reefs sampled at
the peak of bleaching and eight months later. Heat exposure for each reef
was measured as satellite-derived DHW (in °C-weeks).

with simpler morphological characteristics and slower growth rates. By
contrast, on less-bleached reefs the weighted abundances of function-
ally important traits typically showed small gains (Fig. 3f).

In conclusion, our analyses show that acute heat stress from global
warming is a potent driver of a 1,000 km-scale transformation of coral
assemblages, affecting even the most remote and well-protected reefs
within an iconic World Heritage Area. Forecasts of coral bleaching
made continuously by the US National Oceanic and Atmospheric
Administration are accompanied with guidance that a DHW expo-
sure of 4 °C-weeks is expected to cause significant bleaching, and
8 °C-weeks may also result in mortality of corals[15,16,18]. Similarly,
a model for predicting the locations of resilient reefs on the Great
Barrier Reef assumed that coral mortality starts to occur only once
thermal exposure exceeds 6 °C-weeks[19]. However, we show that sub-
stantial mortality occurred on the Great Barrier Reef in 2016 well
below 6 °C-weeks, beginning instead at 3–4 °C-weeks, and with typi-
cal losses exceeding 50% at 4–5 °C-weeks (Fig. 2c). Furthermore, the
threshold that we have identified for the breakdown of assemblage
structure, approximately 6 °C-weeks (Fig. 4), was transgressed in
2016 throughout most of the northern, as well as much of the cen-
tral, region of the Great Barrier Reef (Fig. 1). The prospects for a full
recovery to the pre-bleaching coral assemblages are poor, for several
reasons. First, many of the surviving coral colonies continue to die
slowly even after recovery of their algal symbionts, because they
have lost extensive patches of tissue, are injured and fragmented, and
because corals weakened by bleaching are susceptible to subsequent
outbreaks of disease[20,21]. Second, the replacement of dead corals by
larval recruitment and subsequent colony growth will take at least a
decade even for fast-growing, highly fecund corals, such as species of
Acropora, *Pocillopora*, *Seriatopora* and *Stylophora*[22,23]. The success of
future recruitment will depend on an adequate supply of larvae from
lightly bleached locations, the rapid break down of many millions of
dead coral skeletons to provide a more enduring and stable substrate
for settling larvae and the availability of suitable settlement cues and
conditions for survival of juvenile corals[24]. Third, for longer-lived,
slow-growing species, the trajectory of replacement of dead corals on
heavily damaged reefs will be far more protracted, almost certainly
decades longer than the return-times of future bleaching events. The

recurrence of mass bleaching during the recovery period will be crit-
ical, in view of the global rise in the frequency of bleaching events[4–6].

The 2015–2016 global bleaching event is a watershed for the Great
Barrier Reef, and for many other severely affected reefs elsewhere in
the Indo-Pacific Ocean[4]. Furthermore, the Great Barrier Reef experi-
enced severe bleaching again in early 2017, causing additional exten-
sive damage[25,26]. The most likely scenario, therefore, is that coral reefs
throughout the tropics will continue to degrade over the current cen-
tury until climate change stabilizes[7,27], allowing remnant populations to
reorganize into novel, heat-tolerant reef assemblages. The 2016 marine
heatwave has triggered the initial phase of that transition on the north-
ern, most-pristine region of the Great Barrier Reef (Figs. 1, 4), changing
it forever as the intensity of global warming continues to escalate. The
large-scale loss of functionally diverse corals is a harbinger of further
radical shifts in the condition and dynamics of all ecosystems, rein-
forcing the need for risk assessment of ecosystem collapse[3], especially
if global action on climate change fails to limit warming to 1.5–2 °C
above the pre-industrial base-line.

Online content

Any Methods, including any statements of data availability and Nature Research
reporting summaries, along with any additional references and Source Data files,
are available in the online version of the paper at https://doi.org/10.1038/s41586-
018-0041-2.

1. IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability* (eds Field, C. B. et al.) (Cambridge Univ., Cambridge, 2014).
2. Australian Bureau of Meteorology. 2016 marine heatwave on the Great Barrier Reef. http://www.bom.gov.au/environment/doc/marine-heatwave-2016.pdf (2016).
3. Bland, L. M. et al. Developing a standardized definition of ecosystem collapse for risk assessment. *Front. Ecol. Environ.* **16**, 29–36 (2018).
4. Hughes, T. P. et al. Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science* **359**, 80–83 (2018).
5. Heron, S. F., Maynard, J. A., van Hooidonk, R. & Eakin, C. M. Warming trends and bleaching stress of the World's coral reefs 1985–2012. *Sci. Rep.* **6**, 38402 (2016).
6. Donner, S. D., Rickbeil, G. J. M. & Heron, S. F. A new, high-resolution global mass coral bleaching database. *PLoS ONE* **12**, e0175490 (2017).
7. Hughes, T. P. et al. Coral reefs in the Anthropocene. *Nature* **546**, 82–90 (2017).
8. Baird, A. H. & Marshall, P. A. Mortality, growth and reproduction in scleractinian corals following bleaching on the Great Barrier Reef. *Mar. Ecol. Prog. Ser.* **237**, 133–141 (2002).
9. Baker, A. C., Glynn, P. W. & Riegl, B. Climate change and coral reef bleaching: an ecological assessment of long-term impacts, recovery trends and future outlook. *Estuar. Coast. Shelf Sci.* **80**, 435–471 (2008).
10. Marshall, P. A. & Baird, A. H. Bleaching of corals on the Great Barrier Reef: differential susceptibilities among taxa. *Coral Reefs* **19**, 155–163 (2000).
11. Loya, Y. et al. Coral bleaching: the winners and the losers. *Ecol. Lett.* **4**, 122–131 (2001).
12. Hughes, T. P. et al. Climate change, human impacts, and the resilience of coral reefs. *Science* **301**, 929–933 (2003).
13. Swain, T. D. et al. Coral bleaching response index: a new tool to standardize and compare susceptibility to thermal bleaching. *Glob. Change Biol.* **22**, 2475–2488 (2016).
14. Hughes, T. P. et al. Global warming and recurrent mass bleaching of corals. *Nature* **543**, 373–377 (2017).
15. Eakin, C. M. et al. Caribbean corals in crisis: record thermal stress, bleaching, and mortality in 2005. *PLoS ONE* **5**, e13969 (2010).
16. Liu, G. et al. Reef-scale thermal stress monitoring of coral ecosystems: new 5-km global products from NOAA Coral Reef Watch. *Remote Sens.* **6**, 11579–11606 (2014).
17. Beeden, R. et al. Impacts and recovery from severe tropical cyclone Yasi on the Great Barrier Reef. *PLoS ONE* **10**, e0121272 (2015).
18. Kayanne, H. Validation of degree heating weeks as a coral bleaching index in the northwestern Pacific. *Coral Reefs* **36**, 63–70 (2017).
19. Hock, K. et al. Connectivity and systemic resilience of the Great Barrier Reef. *PLoS Biol.* **15**, e2003355 (2017).
20. Muller, E. M. et al. Bleaching increases likelihood of disease on *Acropora palmata* (Lamarck) in Hawksnest Bay, St. John, US Virgin Islands. *Coral Reefs* **27**, 191–195 (2008).
21. Miller, J. et al. Coral disease following massive bleaching in 2005 causes 60% decline in coral cover on reefs in the US Virgin Islands. *Coral Reefs* **28**, 925–937 (2009).
22. Kayanne, H., Harii, S., Ide, Y. & Akimoto, F. Recovery of coral populations after the 1998 bleaching on Shiraho Reef, in the southern Ryukyus, NW Pacific. *Mar. Ecol. Prog. Ser.* **239**, 93–103 (2002).

23. Gilmour, J. P., Smith, L. D., Heyward, A. J., Baird, A. H. & Pratchett, M. S. Recovery of an isolated coral reef system following severe disturbance. *Science* **340**, 69–71 (2013).
24. Webster, N. S., Soo, R., Cobb, R. & Negri, A. P. Elevated seawater temperature causes a microbial shift on crustose coralline algae with implications for the recruitment of coral larvae. *ISME J.* **5**, 759–770 (2011).
25. Hughes, T. P. & Kerry, J. T. Back-to-back bleaching has now hit two-thirds of the Great Barrier Reef. *The Conversation* https://theconversation.com/back-to-back-bleaching-has-now-hit-two-thirds-of-the-great-barrier-reef-76092 (2017).
26. Great Barrier Reef Marine Park Authority. Final report: 2016 coral bleaching event on the Great Barrier Reef. http://elibrary.gbrmpa.gov.au/jspui/bitstream/11017/3206/1/Final-report-2016-coral-bleaching-GBR.pdf (2017).
27. Hartmann, D. L. et al. in *Climate Change 2013: The Physical Science Basis*. (eds Stocker, T.F. et al.) 159–254 (IPCC, Cambridge Univ. Press, Cambridge, 2013).

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0041-2.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0041-2.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to T.P.H.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Initial mortality and heat stress. We used aerial surveys, conducted in March–April 2016, to measure the geographical extent and severity of bleaching on the Great Barrier Reef, and subsequently converted the bleaching scores into mortality estimates (Fig. 1a) using a calibration curve based on underwater measurements of coral losses (Extended Data Fig. 1). The aerial surveys were conducted throughout the Great Barrier Reef Marine Park and the Torres Strait between Australia and Papua New Guinea, from the coast of Queensland to the outermost reefs, and along the entire Reef from latitudes 9.5 °S to 23.5 °S. Each of 1,156 individual reefs was scored into one of five bleaching categories: 0, less than 1% of corals bleached; 1, 1–10%; 2, 10–30%; 3, 30–60%; 4, more than 60% of corals bleached. The accuracy of the aerial scores was ground-truthed by measuring the extent of bleaching underwater on 104 reefs, also during March–April 2016[14,28].

We assessed underwater the initial mortality of different taxa due to heat stress, at the same time as the aerial surveys, on 63 reefs that spanned the full spectrum of heat exposures and bleaching. On each reef, the extent of bleaching and mortality on individual coral colonies was measured at two sites using five $10 \times 1\,m^2$ belt transects placed on the reef crest at a depth of 2 m. We identified each colony (at the species or genus level) and recorded a categorical bleaching score for each one ($n = 58,414$ colonies): 1, no bleaching; 2, pale; 3, 1–50% bleached; 4, 51–99% bleached; 5, 100% bleached; 6, recently dead. The dead colonies, which had suffered whole-colony mortality, were white with fully intact fine-scale skeletal features, typically still had patches of rotting coral tissue and were experiencing the initial week or two of colonization by filamentous algae, features which distinguished them from corals that had died earlier. The timing of our initial underwater censuses, at the peak of the bleaching in March–April 2016, was critical for identifying corals that were dying directly from heat stress, and for measuring the baseline composition of the assemblages.

Heat stress on the Great Barrier Reef in 2016 was quantified at 5-km resolution, using the NOAA Coral Reef Watch version 3 DHW metric[16]. DHW values are presented in Fig. 1b as a heat map (stretch type: histogram equalize) using inverse distance weighting (power: 2, cell size: 1000, search radius: variable, 100 points) in ArcMap 10.2.1.

Longer term mortality. To measure longer term coral loss (decrease in coral cover after eight months) and its relationship to the level of bleaching and heat exposure, we also conducted detailed before–after assessments of taxon-specific abundances by re-visiting the 63 reefs. We measured abundances in March–April and eight months later at the same locations in October–November, allowing us to compare changes in coral cover for 15 ecologically and taxonomically distinct components of benthic assemblages, on reefs exposed to a broad spectrum of heat stress. These measurements were conducted at the same two geo-referenced sites per reef, on reef crests at a depth of 2 m, using five 10-m long line-intercept transects per site. There were no cyclones or flood events on the Great Barrier Reef during the March–November period (Austral winter) in 2016. Unbleached reefs typically showed small increases in cover due to growth, which we included in the regression analyses. Analysis of change in coral cover was undertaken using the $\log_{10}$-transformed ratio of final to initial cover. To improve readability of Fig. 2 and Extended Data Fig. 1, changes in coral cover are presented as percentages calculated from the log-scale.

We compared the initial and final composition of corals using a non-metric multi-dimensional scaling (nMDS) analysis based on a Bray–Curtis similarity matrix of square-root transformed data, and quantified the shift over time using the Euclidean distance between before–after assemblages at each location. We then estimated the relationship between the shift in composition at each reef versus the level of heat exposure experienced there (Fig. 4). To include all species, the majority of which are too rare to analyse individually, we pooled them into 15 ecologically cohesive groups depending on their morphology, life history and taxonomy. Three of the fifteen groups are ubiquitous species or species complexes: *Pocillopora damicornis*, *Seriatopora hystrix* and *Stylophora pistillata*. In each of the multi-species groups, the dominant species or genera on reef crests were: other *Acropora* (*A. gemmifera*, *A. humilis*, *A. loripes*, *A. nasuta*, *A. secale*, *A. tenuis* and *A. valida*); favids (that is, species and genera from the formerly recognized family Faviidae: *Cyphastrea*, *Favia*, *Favites*, *Goniastrea*, *Leptastrea*, *Montastrea* and *Platygyra*); Mussidae (*Lobophyllia* and *Symphyllia*); *Isopora* (*I. palifera* and *I. cuneata*); other *Pocillopora* (*P. meandrina* and *P. verrucosa*); other sessile animals (sponges, tunicates, molluscs); *Porites* (*P. annae* and *P. lobata*); *Montipora* (*M. foliosa*, *M. grisea*, *M. hispida*, *M. montasteriata* and *M. tuberculosa*); staghorn *Acropora* (*A. florida*, *A. intermedia*, *A. microphthalma*, *A. muricata* and *A. robusta*); soft corals (alcyonaceans and zooanthids); tabular *Acropora* (*A. cythera*, *A. hyacinthus* and *A. anthocercis*).

We calculated longer term mortality for all species combined at the scale of the entire Great Barrier Reef in three ways, all of which yielded consistent results. The first approach, which provided the best spatial resolution (Fig. 1a), was based on a comparison of the observed loss of total coral cover on 63 reefs that extend along the entire Great Barrier Reef measured underwater between March and November, with aerial bleaching scores of the same locations in March–April (Extended Data

Fig. 1). This calibration allowed us to convert the aerial scores of bleaching that we recorded for 1,156 reefs into mortality estimates for each of the five aerial score categories, and to map the geographic footprint of losses of corals throughout the Great Barrier Reef (Fig. 1a). The spatial patterns of coral decline (Fig. 1a) are presented as a heat map of the calibrated scores (stretch type: histogram equalize) using inverse distance weighting (power: 2, cell size: 1000, search radius: variable, 100 points) in ArcMap 10.2.1.

The second methodology for estimating large-scale mortality is independent of aerial surveys of bleaching, and based on the loss of total coral cover on 110 reefs (Extended Data Fig. 2), including the 63 reefs that were re-censused for change in composition. The median cover on these reefs declined between March and November from 34% to 20% (Extended Data Fig. 3). For method two, the observed loss of coral cover was averaged for replicate reefs surveyed within each of eight sectors of the Great Barrier Reef Marine Park and the Torres Strait), corrected for differences in reef area for each sector based on GIS data provided by the Great Barrier Reef Marine Park Authority, and then summed to calculate the total loss. For method three, we used the fitted relationship between satellite-derived DHW and observed change in cover (Fig. 2c) to score the losses or gains on all 3,863 individual reefs comprising the Great Barrier Reef, and averaged the total. These two alternative approaches for estimating large-scale loss of cover, both based on before–after underwater surveys (Extended Data Figs. 2 and 3) yielded consistent results with Fig. 1a—a 29.0% and 27.7% decline, respectively, after eight months.

Differential mortality among coral taxa. To estimate how exposure to heat (measured as DHW) affects loss of cover differentially among taxa, we used a linear mixed effects model. The fixed effect was DHW and we allowed for a random effect of taxonomic grouping on both the intercept and slope of the relationship between coral cover change and DHW. We excluded from the analysis observations with zero initial coral cover of a particular taxonomic group. Change in coral cover was transformed before analysis by calculating the $\log\left(\frac{C_f + \epsilon}{C_i + \epsilon}\right)$ where C_f and C_i were the final and initial coral cover, respectively, and ϵ was the minimum observed value of coral cover. The estimated random effect on intercepts was approximately zero, so we eliminated it from our final model. Thus, in the final model, there was a common intercept, but differences between taxa in sensitivity to DHW (that is, there was a random effect of taxonomic group on the slope). To illustrate these differences, Extended Data Fig. 5 plots the estimated slope of the coral cover response variable for each taxon versus DHW as the overall mean effect of DHW plus the taxon-specific random effect. Conditional standard errors plotted in Extended Data Fig. 5 are the standard errors on each random effect.

Shifts in functional traits. To calculate how differential mortality affected the mix of traits in the coral assemblages, we scored eight traits for 12 of the 15 functional groupings (excluding soft corals, other Scleractinia, and other sessile fauna, Extended Data Tables 1, 2). We chose traits that are likely to influence ecosystem functions. For example, corals with fast growth rates and high skeletal density strongly influence calcification, colony shape affects photosynthesis and the provision of three-dimensional habitat, and the size of corallites is a measure of heterotrophy. The traits were scored using the Coral Trait Database[29], with the exception of colony size, which we measured directly for each group on reef crests using the geometric mean of intercept lengths for each taxon from our initial transects. For multi-species groups, the traits were generally identical for all species, except for *Montipora* and *Porites*, for which we used the mean score across the reef crest species that we encountered. To measure the depletion of traits based on changes in absolute abundances between March and November (Fig. 3e, f), we used a community weighted mean (CWM) analysis of each trait:

$$\mathrm{CWM} = \sum_{i=1}^{n} a_i \mathrm{trait}_i$$

where a_i is the abundance of coral taxa i and trait_i is the trait value of coral taxa i. This metric provides a trait value for each reef weighted by the total abundance of each taxa. To visualize the overall shift in functional composition, we used a nMDS analysis based on a Bray–Curtis dissimilarity matrix of square-root transformed data for each trait community weighted mean, creating a multi-dimensional trait space in which reefs are positioned according to the value and abundance of critical traits.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. All heat exposure data used in this study are publicly available from the US National Oceanic and Atmospheric Administration. Source data for coral bleaching, mortality and abundances are available online at the Tropical Data Hub: https://doi.org/10.4225/28/5a725ee7548a7.

28. Hughes, T. P., Kerry, J. T. & Simpson, T. Large-scale bleaching of corals on the Great Barrier Reef. *Ecology* **99**, 501 (2017).
29. Madin, J. S. et al. The Coral Trait Database, a curated database of trait information for coral species from the global oceans. *Sci. Data* **3**, 160017 (2016).
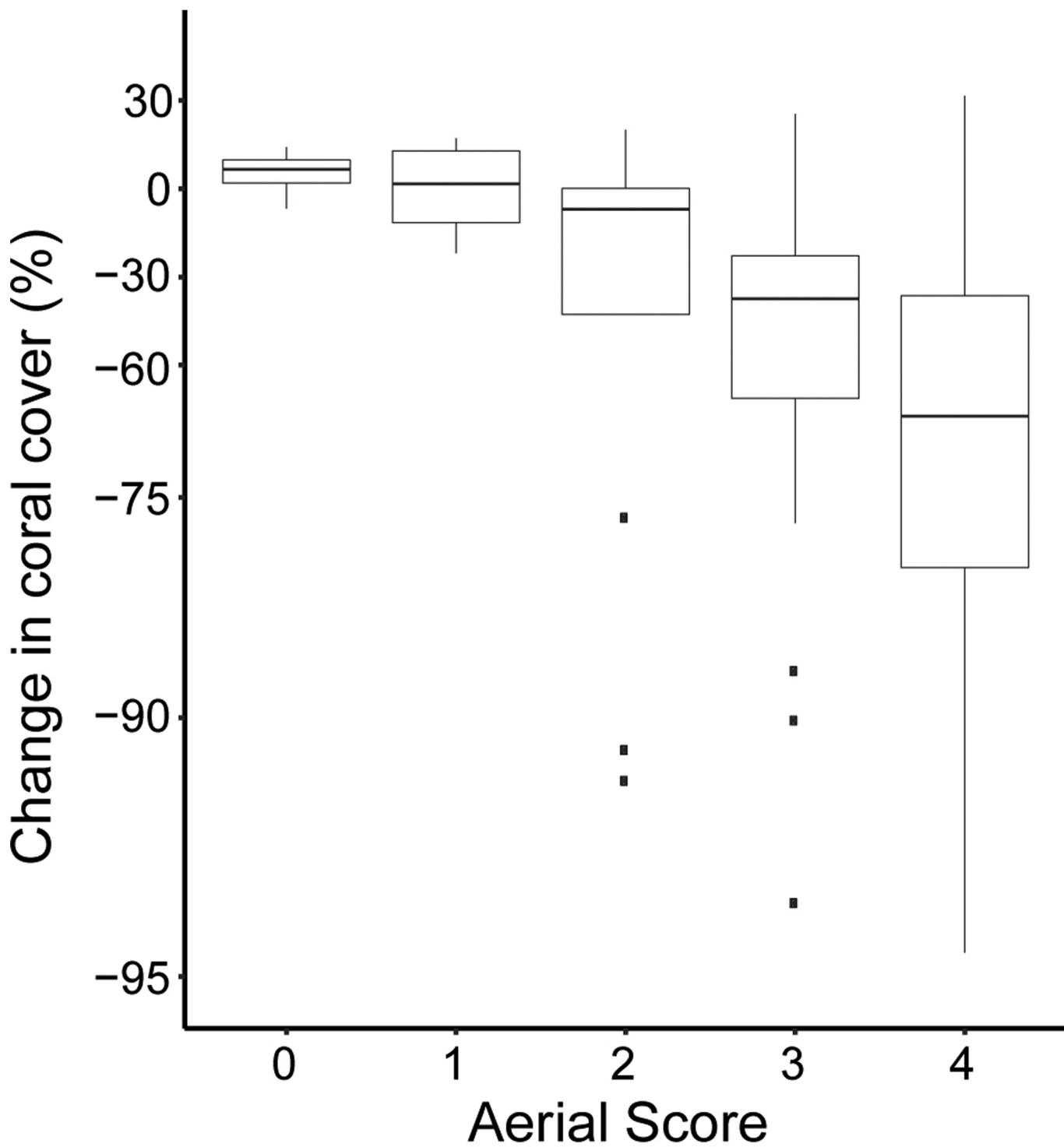
Extended Data Fig. 1 | Relationship between aerial bleaching scores and change in coral cover. Aerial scores of bleaching on the *x* axis are: 0 (< 1% of colonies bleached), 1 (1–10%), 2 (10–30%), 3 (30–60%) and 4 (60–100%). Change in coral cover on the *y* axis was measured in situ between March and November 2016 on 98 reefs that were also scored from the air. Box plots are shown for each aerial category, showing median values (horizontal lines), boxes for values in the 25th–75th percentiles, vertical lines for values less than the 25th percentile and greater than the 75th, and data points for outliers. Medians were used when calibrating change in cover for each aerial category (see Fig. 1a).
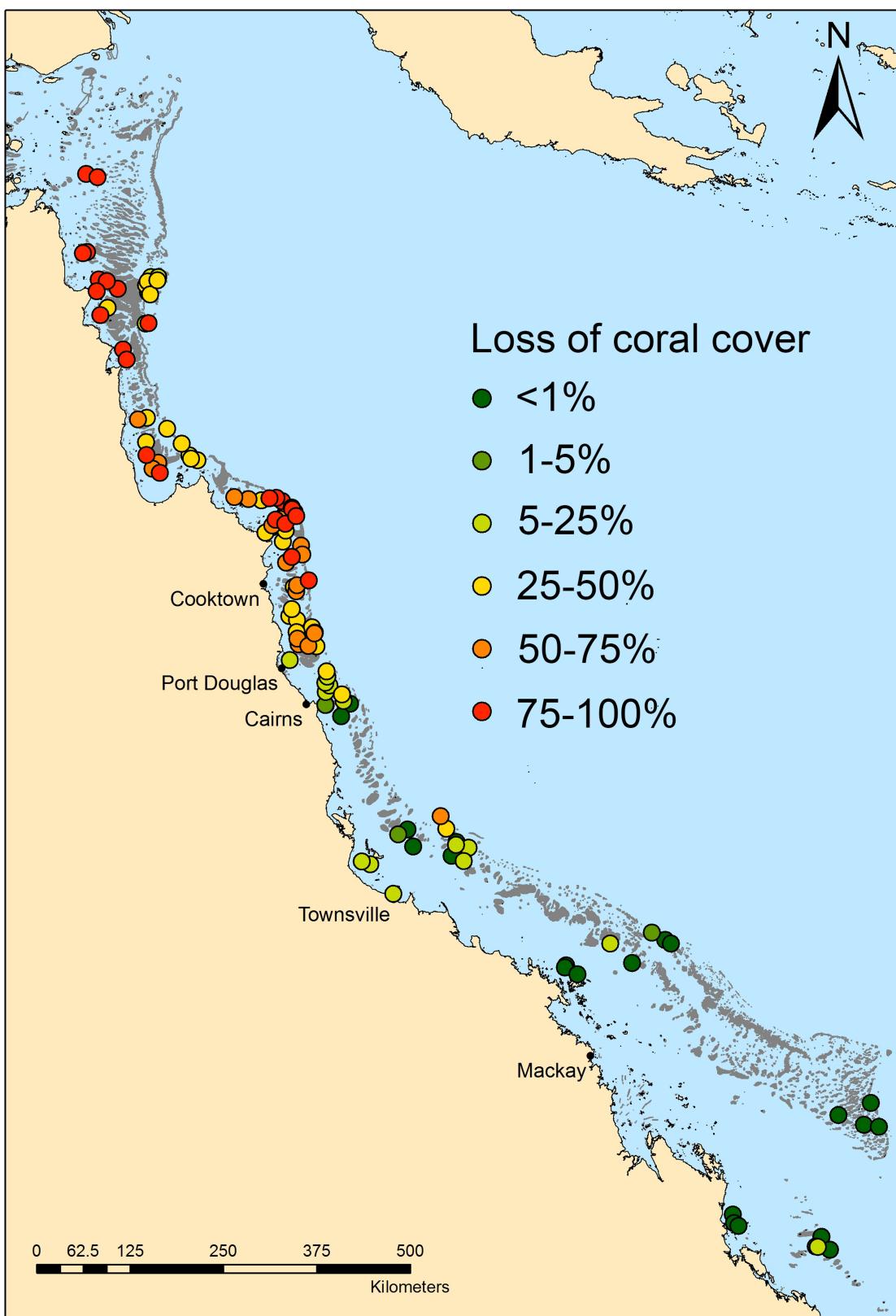
Extended Data Fig. 2 | Loss of coral cover along the Great Barrier Reef in 2016. Losses, measured on 110 reefs between March and November 2016, range from 0 (dark green) to 100% (1–5% (green), 5–25% (light green), 25–50% (yellow), 50–75% (orange) and 75–100% (red)). Map template is provided by Geoscience Australia (Commonwealth of Australia (Geoscience Australia) 2018).
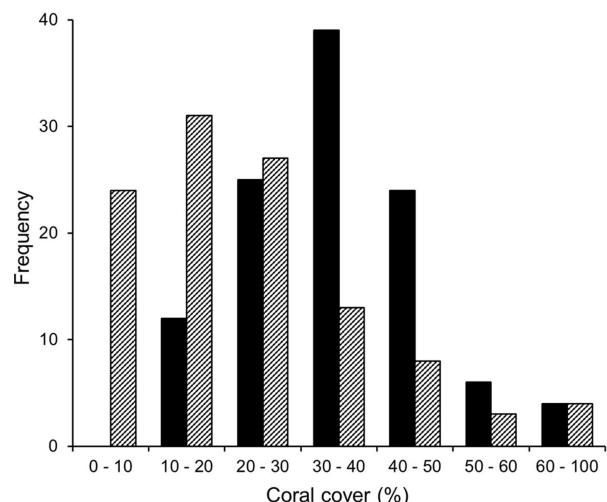
Extended Data Fig. 3 | Shifts in coral cover following coral bleaching.
The frequency distribution of coral cover on 110 reefs, measured in
March 2016 (solid bars) and again in November 2016 (hashed bars). Reef
locations are shown in Extended Data Fig. 2.
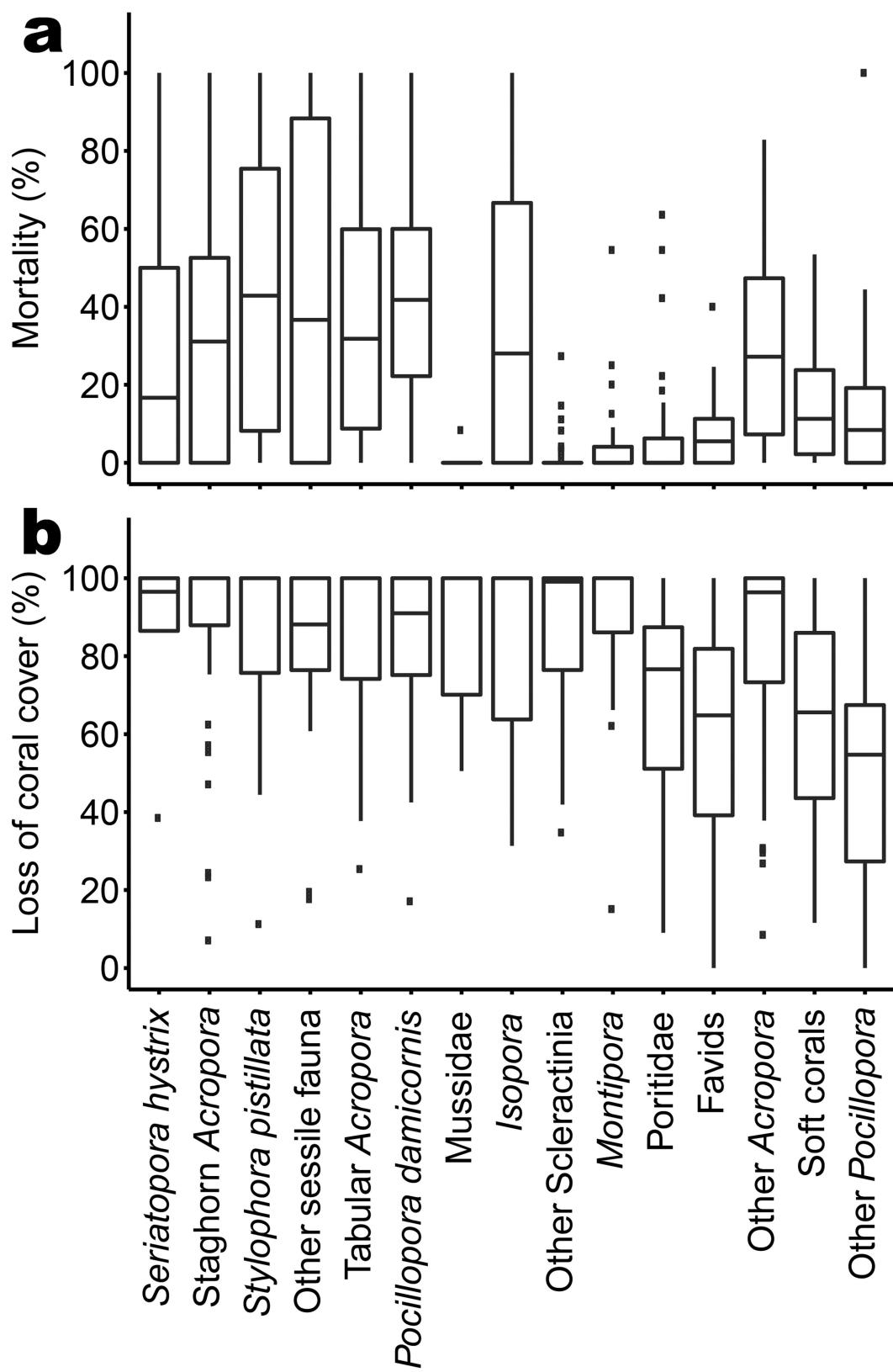
Extended Data Fig. 4 | Mortality rates differ among coral taxa. Box plots are shown for each taxon, showing median mortality (horizontal lines), boxes for the middle two quartiles, vertical lines for the first and fourth quartiles, and data points for outliers. **a**, The initial mortality of corals recorded on belt transects on 43 reefs with > 60% bleaching. **b**, Longer term loss of cover for taxonomic categories recorded between March and November 2016 on the 43 remeasured reefs with > 60% bleaching. Taxa in **a** and **b** are plotted in rank order along the *x* axis, from highest to lowest decreases in mean cover between March and November 2016.
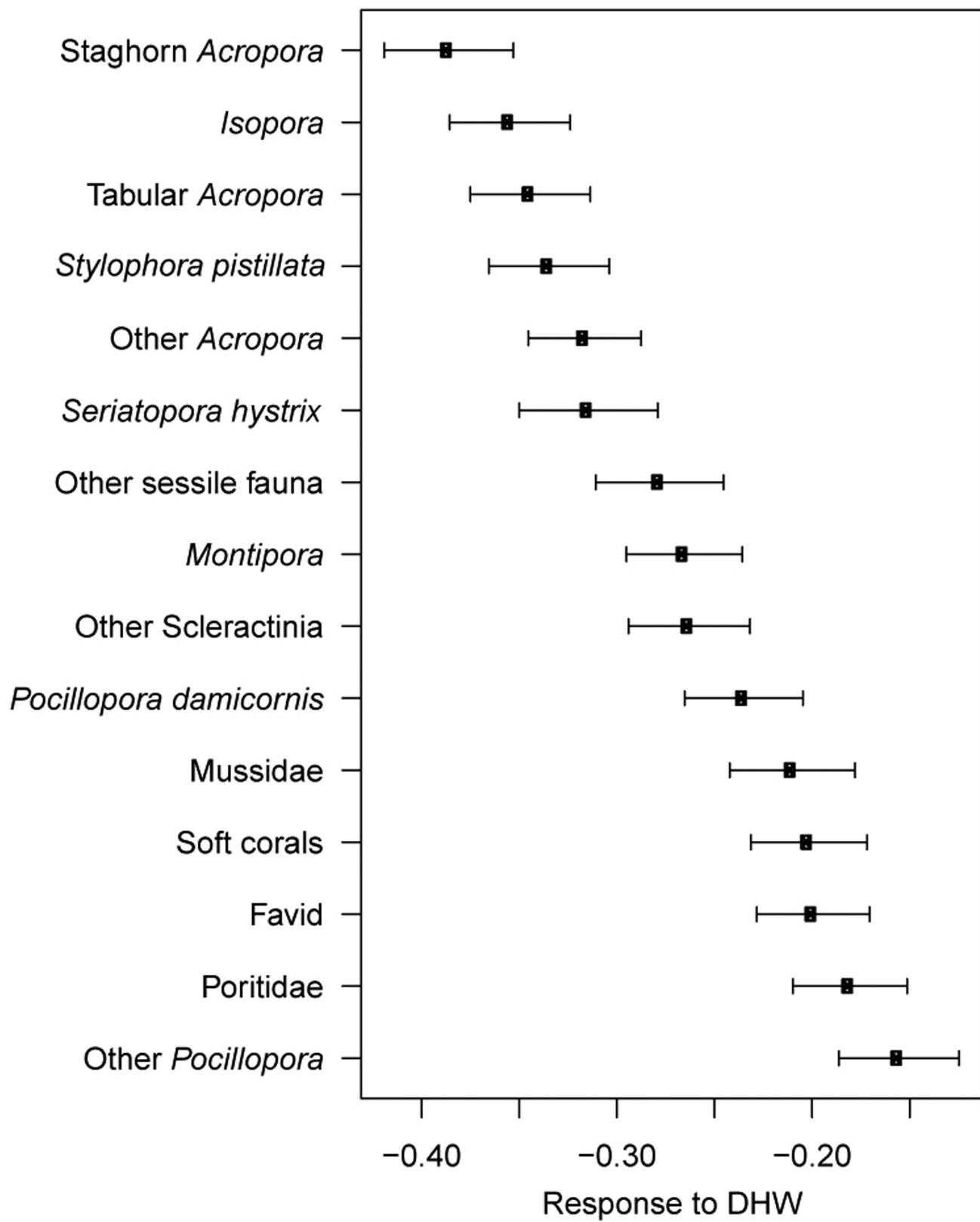
Extended Data Fig. 5 | Differential sensitivity of coral taxa to temperature stress. Sensitivity is estimated from the loss of cover on 63 reefs for different groups of corals between March and November 2016, as a function of heat exposure (DHW). The horizontal axis is the slope of the relationship between the log-ratio of final and initial coral cover (response variable) and DHW (explanatory variable). Values plotted for each taxonomic grouping (ordered from most sensitive to least sensitive) are random effects estimates, with conditional standard errors.
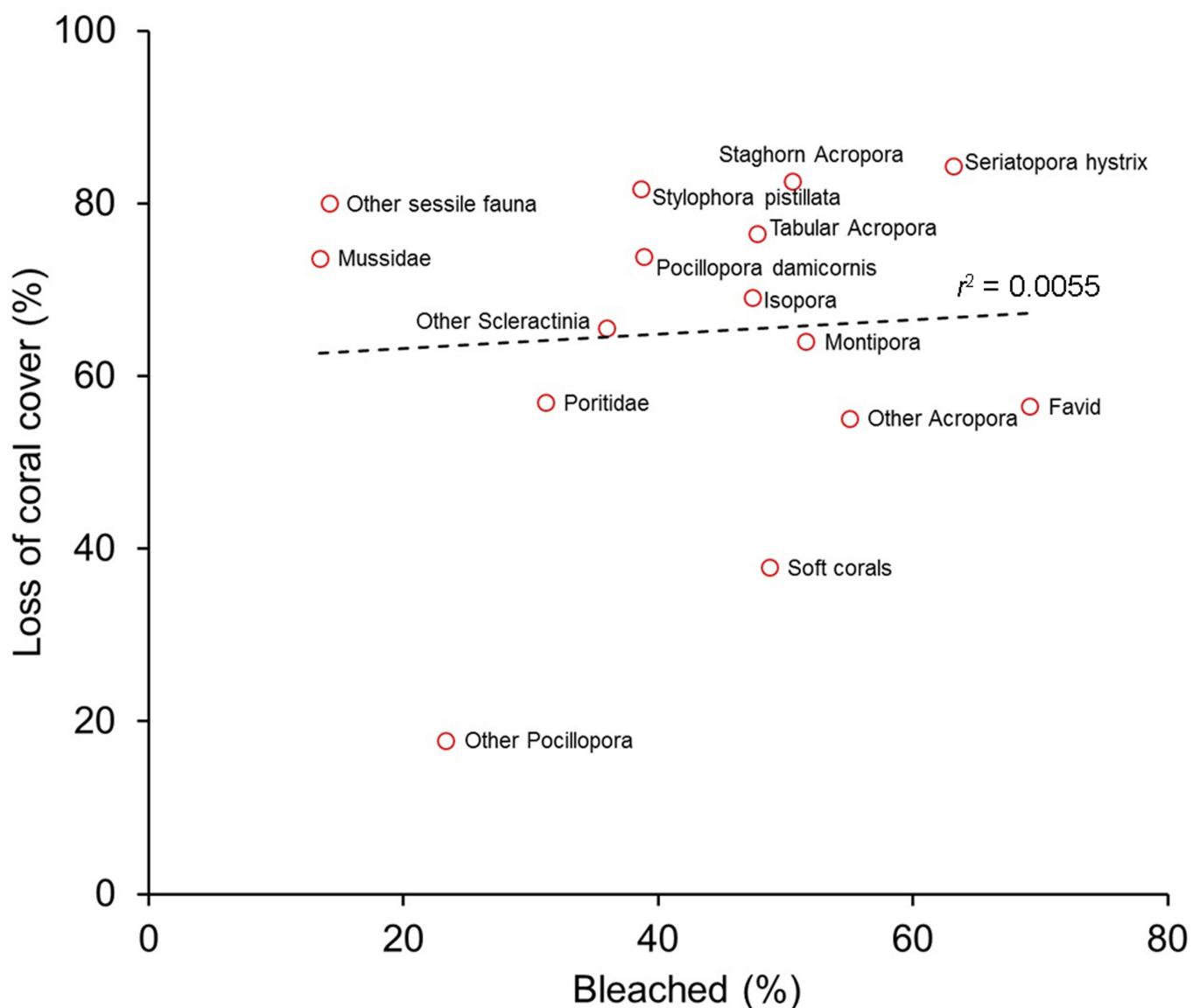
Extended Data Fig. 6 | Bleaching extent is unrelated to mortality. The regression shows the relationship between the levels of bleaching by individual coral taxa on severely bleached reefs (where > 60% of all colonies were affected, $n = 43$ reefs), and their subsequent loss of cover eight months later. The non-significant correlation indicates that the winners–losers spectrum of bleaching among taxa is a poor predictor of which ones ultimately die.

Extended Data Table 1 | Eight traits of coral species and their key functional roles

Trait	Trait scores	Reef function
Growth rate	In mm/year: 0-10 (1), 10-20 (2), 20-40 (3), 40-60 (4), >60 (5).	Carbonate framework accretion; reef regeneration
Skeletal density	In g/cm^3: <1 (1), 1-1.4 (2), 1.4-1.7 (3), 1.7-2 (4), >2 (5)	Carbonate framework accretion
Corallite width	In mm: <1 (1), 1-2 (2), 2-5 (3), 5-15 (4) ; <15 (5)	Filter feeding; nutrient capture
Interstitial space size	(1-5) Based on morphological categories.	Habitat provision
Colony height	(1-5) Based on morphological categories.	Carbonate framework accretion; habitat provision
Surface area to volume ratio	(1-5) Based on morphological categories	Primary productivity; nutrient cycling
Colony size	Rank (1-12) measured from reef crest transects	Carbonate framework accretion; habitat provision
Reproductive mode	Brooders (1), Mixed (2), Spawners (3)	Reef connectivity and regeneration

Extended Data Table 2 | Trait scores for each of 12 groups of corals

Taxon	Corallite size	Growth rate	Colony size	Skeletal density	Colony height	Tissue area	Interstitial space size	Reproductive mode
Bushy *Acropora*	2	3	7	3	3	5	3	Spawner
Favids	4	1	4	3	2	1	1	Spawner
Isopora	2	2	10	3	2	2	1	Brooder
Montipora	2	3	9	5	1	1	1	Spawner
Mussidae	5	1	3	2	2	1	1	Spawner
Other *Pocillopora*	1	3	8	3	3	4	3	Spawner
Pocillopora damicornis	1	3	2	4	2	4	3	Brooder
Poritidae	2	2	6	2	4	1	1	Mix
Seriatopora hystrix	1	3	1	5	2	3	3	Brooder
Staghorn *Acropora*	2	5	11	4	5	3	5	Spawner
Stylophora pistillata	2	3	5	4	2	3	3	Brooder
Tabular *Acropora*	2	4	12	4	3	5	5	Spawner

Spawners release eggs and sperm that fertilize externally, whereas brooders release internally fertilized planulae larvae.

LETTER

Vertically migrating swimmers generate aggregation–scale eddies in a stratified column

Isabel A. Houghton[1], Jeffrey R. Koseff[1], Stephen G. Monismith[1] & John O. Dabiri[1,2]*

Biologically generated turbulence has been proposed as an important contributor to nutrient transport and ocean mixing[1–3]. However, to produce non-negligible transport and mixing, such turbulence must produce eddies at scales comparable to the length scales of stratification in the ocean. It has previously been argued that biologically generated turbulence is limited to the scale of the individual animals involved[4], which would make turbulence created by highly abundant centimetre-scale zooplankton such as krill irrelevant to ocean mixing. Their small size notwithstanding, zooplankton form dense aggregations tens of metres in vertical extent as they undergo diurnal vertical migration over hundreds of metres[3,5,6]. This behaviour potentially introduces additional length scales—such as the scale of the aggregation—that are of relevance to animal interactions with the surrounding water column. Here we show that the collective vertical migration of centimetre-scale swimmers—as represented by the brine shrimp Artemia salina— generates aggregation-scale eddies that mix a stable density stratification, resulting in an effective turbulent diffusivity up to three orders of magnitude larger than the molecular diffusivity of salt. These observed large-scale mixing eddies are the result of flow in the wakes of the individual organisms coalescing to form a large-scale downward jet during upward swimming, even in the presence of a strong density stratification relative to typical values observed in the ocean. The results illustrate the potential for marine zooplankton to considerably alter the physical and biogeochemical structure of the water column, with potentially widespread effects owing to their high abundance in climatically important regions of the ocean[7].

Biologically driven macro-scale flow has been observed in systems ranging from the bio-convective patterns generated by dense bacterial suspensions[8,9] to large-scale induced drift by individually swimming jellyfish[2]. At intermediate length scales, marine zooplankton—which are individually approximately a centimetre in scale—undergo diurnal vertical migration in dense swarms over distances of hundreds of metres[5–7]. Dynamic feedbacks between the zooplankton propulsion and the surrounding flow have the potential to affect the physical and biogeochemical structure of the water column, and they are particularly relevant owing to the presence of diurnal vertical migration in climatically important regions such as the Southern Ocean[7]. Although turbulence created by an individually swimming zooplankton is limited by the length scale of the animal and is therefore primarily dissipated as heat rather than contributing to turbulent mixing[4], the collective migration introduces a new length scale in the form of the vertical height of the swarm, which has previously been observed using acoustic backscatter to span tens of metres[3,5,6].

Although large-scale effects on nutrient distributions, particularly carbon and oxygen, due to the presence of diurnal vertical migration have been reported[7,10,11], direct observations of the fluid dynamics surrounding zooplankton swarms have been limited owing to the difficulty of predicting the precise time and location of a migration event. Building on a previously developed method[12], we conducted laboratory experiments using the representative centimetre-scale swimmers *A. salina* in two stably stratified tank facilities. A 1.2-m-tall tank was used to measure irreversible mixing of the density stratification, and a 2-m-tall tank was used to implement several flow visualization techniques over multiple length scales. In both tanks the strong phototactic response of *A. salina* was leveraged to induce coordinated migrations over the vertical extent of the tank, simulating zooplankton migration through a pycnocline. The 1.2-m tank used an array of focused LEDs with blue filters to form a 10-cm-wide vertical column of light that triggered vertical migration (Fig. 1a). The 2-m tank used a blue laser to form a narrower 5-cm-wide column of light (Fig. 1b). Both light stimuli triggered a similar animal response (for example, in terms of swimming speed and inter-animal spacing), indicating that the migration dynamics were not sensitive to the specific form of visual stimulus.

In experiments measuring irreversible mixing, a stable two-layer stratification was established by using variable-salinity water pumped into the tank in two layers. The average buoyancy frequency across the interface between the two layers (N_{int}) is given by $N_{\text{int}} = \sqrt{-\frac{g}{\rho_0}\frac{\Delta\rho}{\Delta z}}$ in which ρ_0 is the reference density, $\Delta\rho$ is density difference between the top and bottom layer, Δz is the initial interface thickness of 0.2 m used in all cases, and g is gravity. In these experiments, the N_{int} ranged from 0.04 to 0.13 s^{-1}. After stratifying the water column, animals were introduced at a tank-averaged abundance of between 46,000 ± 5,000 animals per m^3 and 138,000 ± 5,000 animals per m^3 in the 1.2-m tank, or 20,000 ± 5,000 animals per m^3 in the 2-m tank, and allowed to acclimatize. Experimental animal abundance was comparable to that observed for a wide variety of ocean zooplankton, which range in abundance[6,13,14] from 10,000 to 70,000 animals per m^3, although local animal densities can be underestimated by orders of magnitude in these types of observations owing to the high spatial heterogeneity in swarms[3,15].

Animals were initially gathered at the bottom of the tank by attracting them with a green LED array introduced to the tank from below. *A. salina* exhibits slight negative buoyancy, similar to most ocean zooplankton, and therefore the animals were minimally active at the bottom of the tank. Initial density profile measurements were then obtained with a vertically traversing density probe located 20 cm laterally from the centre of the LED array (Fig. 1c). Mixing during the filling process smoothed the two-layer stratification interface to produce an error function-like initial density profile (Fig. 2a, black curve).

Animal migration was induced with a blue LED array oriented vertically downward from above the tank, forming a column of light 10 cm in diameter, activated after the deactivation of the green LED array at the bottom of the tank. A strong animal response towards the light occurred, resulting in an upward migration with swimming velocities up to 1 cm s^{-1} (Fig. 1a; see Supplementary Video 1). The distribution of animal reaction times resulted in a vertical spread of animals over the extent of the tank following the activation of the top LED array. After ten minutes, the top blue LED array was deactivated and the bottom green LED array was activated to return the animals to the bottom of the tank. After ten minutes with the animals at the bottom, the bottom

[1]Bob and Norma Street Environmental Fluid Mechanics Laboratory, Civil and Environmental Engineering, Stanford University, Stanford, CA, USA. [2]Mechanical Engineering, Stanford University, Stanford, CA, USA. *e-mail: jodabiri@stanford.edu
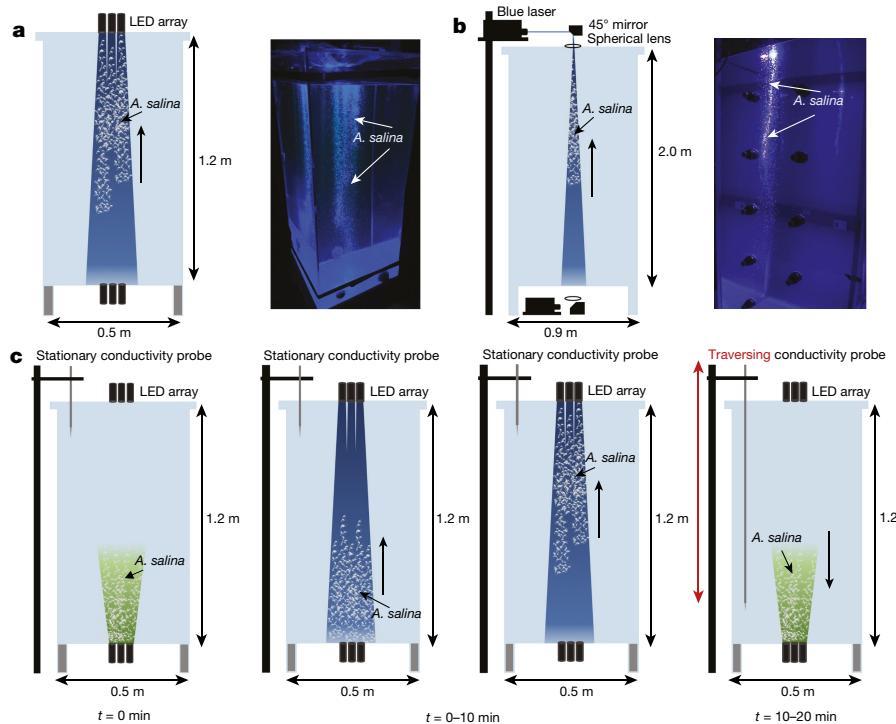
Fig. 1 | Laboratory experiment for controllable vertical migrations. Schematics and images of measurement facilities during experiments are shown. **a**, A $0.5 \times 0.5 \times 1.2$-m^3 tank with a two-layer salt stratification was used with an array of 12 focused LEDs with blue filters introduced from the top of the tank, and LEDs with green filters introduced from below for control of the migration. **b**, A $0.9 \times 0.9 \times 2.0$-m^3 tank with a linear salt stratification was designed with a blue laser above and green laser below for control of the migration. **c**, Schematic of experimental protocol for measuring irreversible mixing. The 20-min cycle illustrated was repeated six times in total to give a 120-min experiment. A density profile was obtained during each downward migration period. Repeated density profiles were obtained at the end of the experiment to confirm consistency of the density profiles, which indicated that the tank was quiescent and horizontally homogeneous.

green LED array was deactivated and the top blue LED array was once again activated to repeat the up–down migration cycle (Fig. 1c). This cycle was repeated six times in each full experiment lasting 120 min. The duration of the full experiment was selected to match oceanic diurnal vertical migration with a representative swimming speed of $1\,\mathrm{cm\,s}^{-1}$ and a representative aggregation with a vertical extent[5] of 50–100 m, leading to a time of approximately 80–170 min for the entire swarm to pass through a fixed depth in the water column that is initially above the aggregation. In the laboratory experiments, the water column was perturbed by the same group of animals multiple times during the 120 min, whereas in an oceanic diurnal vertical migration a given parcel of water is sequentially perturbed by multiple, distinct animals over the 120-min period. Density profiles were obtained each of the six times during the experiment when the animals were at the bottom of the tank, as well as at the end of the experiment. Measured mass changes between initial and final density profiles were limited to ±6 p.p.m. of total water column mass, attributable to minor probe noise and drift.

The final density profiles after the 120-min experiment exhibited irreversible mixing of the initial density profile. This was indicated by a smoothing of the initial error function density profile; that is, the upper half of the density profile shifted towards higher density and the lower half shifted towards lower density (Fig. 2). Vertically asymmetric mixing occurred with more mixing above the pycnocline than below.

A depth-dependent effective turbulent diffusivity ($\kappa_{\mathrm{eff}}(z)$) was used to quantify the effect of swimmer-induced mixing relative to the molecular diffusion of salt (Fig. 3). This empirical effective diffusivity was determined on the basis of a conservative (that is, lower mixing) best-fit numerically integrated density profile to approximate the final experimental profile (see Methods). This conservative estimate yielded effective diffusivities three orders of magnitude larger than the molecular diffusivity of salt.

Flow visualization was used to investigate the processes leading to the observed irreversible mixing. At the scale of the individual swimmer,
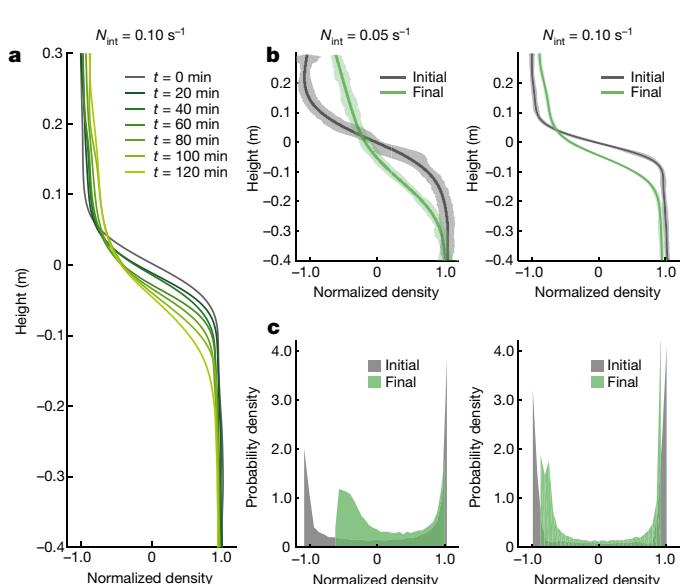
Fig. 2 | Density profile temporal evolution. a, Density profiles for $N_{\mathrm{int}} = 0.10\,\mathrm{s}^{-1}$ taken every 20 min during a 120-min experiment. **b**, Initial and final density profiles for $N_{\mathrm{int}} = 0.05\,\mathrm{s}^{-1}$ (left) and $N_{\mathrm{int}} = 0.10\,\mathrm{s}^{-1}$ (right). Shading corresponds to the average scatter of data due to probe measurement uncertainty, with higher relative uncertainty for the lower density difference (left). Higher density stratification results in more asymmetric mixing (right). Normalized density profiles are presented (see Supplementary Methods). **c**, Probability density plots of normalized density corresponding to $N_{\mathrm{int}} = 0.05\,\mathrm{s}^{-1}$ (left) and $N_{\mathrm{int}} = 0.10\,\mathrm{s}^{-1}$ (right). Density values shift away from extreme values, indicative of irreversible mixing. Measured mass changes between initial and final density profiles were limited to ±6 p.p.m. of total water column mass, attributable to minor probe noise and drift.
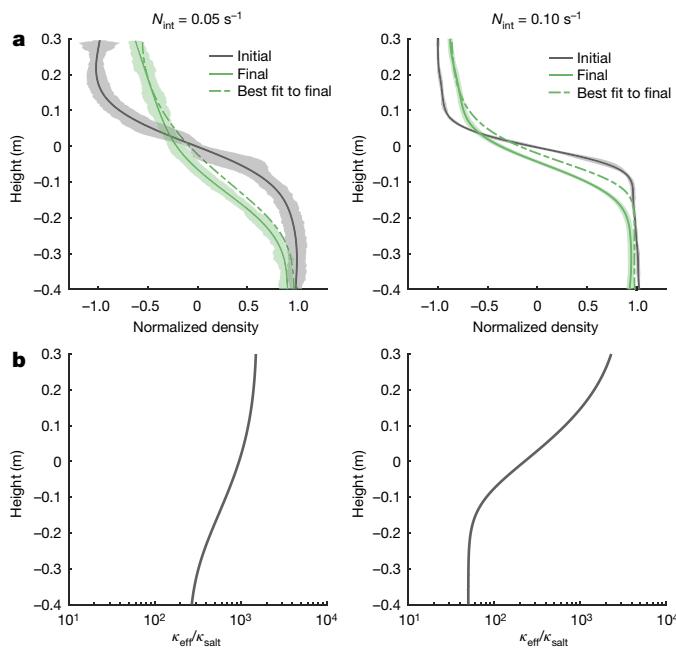
Fig. 3 | Effective diffusivity due to vertical migration. a, Initial (black) and final (solid green) density profiles for $N_{int} = 0.05\,s^{-1}$ (left) and $N_{int} = 0.10\,s^{-1}$ (right). An empirical estimate of effective diffusivity is used for a numerically calculated, best-fit density profile (dashed green; see Supplementary Methods). **b**, Ratio of effective diffusivity to salt molecular diffusivity ($\kappa_{eff}/\kappa_{salt}$) as a function of height for each experiment above. The conservative nature of these estimates is indicated by the smaller shift in the best-fit profile (dashed green curves in **a**) from the initial profiles (black) compared to the final measured profiles (solid green).

A. salina propulsion produces a strong rearward jet (Fig. 4a; see Supplementary Video 2), similar to observed flow fields of zooplankton such as Pacific and Antarctic krill[16]. In the collective migration, a large-scale downward jet developed within the region of upward animal migration owing to the repeated downward fluid displacement by each individual animal propelling itself upward (Fig. 4b–g). Similar to oceanic aggregations[17], downward swimming required less active propulsion owing to the aforementioned negative animal buoyancy, resulting in no aggregation-scale jet during downward migration. This unidirectional formation of a large-scale downward jet during the upward migration resulted in density profiles similar to other asymmetric mixing processes such as horizontal shear at a density interface[18].

Animal proximity in the migration was primarily governed by endogenous animal behaviour, regardless of the light stimulus type and configuration used. The resultant animal spacing was less than the extent of the hydrodynamic signature from an individual animal, similar to observations of Antarctic krill[16]. Owing to this animal proximity, subsequent migrating animals repeatedly interacted with already displaced fluid parcels, propelling fluid further downward and ultimately producing the aggregation-scale rearward jet, characterized by a fluid velocity of $1-2\,cm\,s^{-1}$. Outside the organized animal migration, eddying motions at scales an order of magnitude larger than the individual animal were observed (see Supplementary Videos 3–6), similar to flows observed around negatively buoyant plumes[19]. The large-scale eddy motion within the stable density stratification increased the surface area and concentration gradients between the high and low density fluid, enhancing diffusivity and leading to irreversible mixing of the salt concentration at a rate up to three orders of magnitude larger than molecular action alone. For a temperature-stratified water column, irreversible mixing of a thermal stratification is likely to be higher than the mixing measured in the salinity stratification owing to the faster action of thermal diffusivity compared to salt diffusivity[18,20–22].
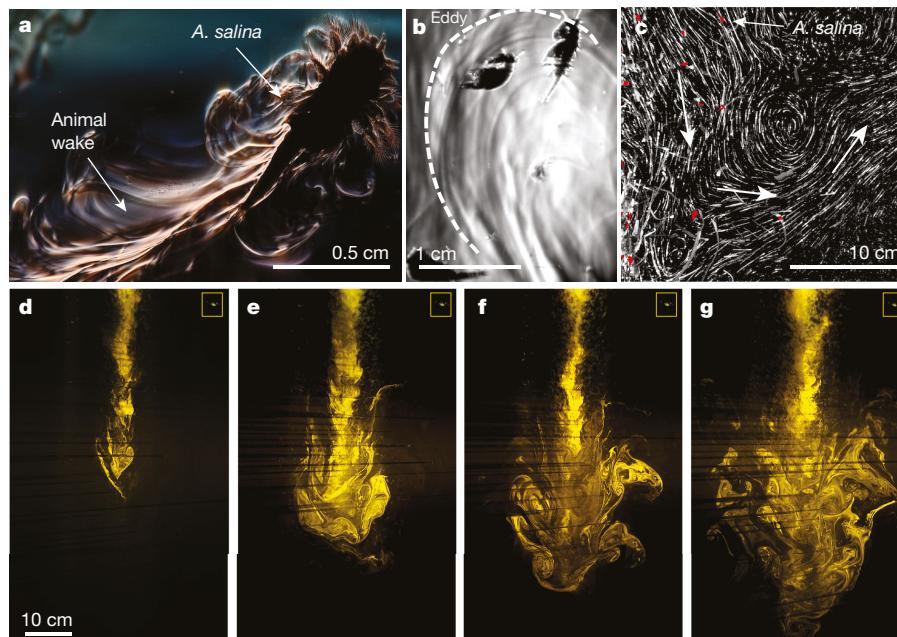
Fig. 4 | Flow visualization from animal to aggregation scales. a, Single animal schlieren image, conducted with the assistance and facilities of R. Strickler, showing fluid motion in the wake of a single animal (see Supplementary Video 2). **b**, Intermediate-scale schlieren imaging with a 5-cm field of view centred 8 cm to the left of the migration, showing the laterally propagating eddies perturbing the stable background density (see Supplementary Video 3). **c**, Pathlines of 10-μm neutrally buoyant particles illustrate the fluid motion to the right of the migration, including the downward jet proximate to the aggregation and eddy motion on the periphery. Individual swimmers are overlaid in red. White arrows show flow direction (see Supplementary Video 5). **d–g**, Planar cross-section of laser-induced fluorescence of a tracer dye propelled downward through the extent of the migration at $t = 54\,s$ (**d**), $122\,s$ (**e**), $185\,s$ (**f**) and $292\,s$ (**g**). Surface fluid is propelled vertically downward more than 50 cm through a stable stratification, with large-scale flow structures entraining fluid proximate to the downward jet. Animals in the laser sheet cast horizontal shadows through the illuminated dye. Individual animal size is highlighted in the inset in the upper right hand corner. All imaging methods are described in detail in the Supplementary Methods.

Laboratory experiments in both the 1.2-m-tall and 2-m-tall tank consistently showed that the downward jet propagated throughout the full vertical extent of the aggregation, that is, wherever animals were present to repeatedly propel fluid downward, even within a strong background stratification relative to typical values observed in the ocean. For an oceanic aggregation of zooplankton, flow developing over this extent would correspond to a vertical span of tens of metres. Because diurnal vertical migrations often extend from the surface to depths of 600 m or deeper[7], the diurnal passage of zooplankton could potentially induce large-scale vertical motion and mixing in regions of high nutrient, microbial and density variability. Such transport and mixing would have important feedbacks for local growth and productivity, thereby affecting regional marine biogeochemistry[23–25].

Moreover, the irreversible mixing observed due to the passage of an aggregation of swimmers has potential implications for physical feedbacks with the density stratification, for example, by changing local dynamics and driving regional flow. The global presence of diurnal vertical migration[7], the substantial horizontal and vertical span of aggregations[6], and the frequent occurrence of migrations presents a mechanism for non-negligible mixing across stabilizing gradients, with numerous potential effects on the physical and biogeochemical structure of the ocean. It remains to observe these mechanisms in situ; however, recent reanalyses of echosounder and acoustic Doppler current profiler measurements suggest that the signature of biogenic ocean mixing is ripe for discovery[17].

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0044-z.

1. Dewar, W. K. et al. Does the marine biosphere mix the ocean? *J. Mar. Res.* **64**, 541–561 (2006).
2. Katija, K. & Dabiri, J. O. A viscosity-enhanced mechanism for biogenic ocean mixing. *Nature* **460**, 624–626 (2009).
3. Huntley, M. E. & Zhou, M. Influence of animals on turbulence in the sea. *Mar. Ecol. Prog. Ser.* **273**, 65–79 (2004).
4. Visser, A. W. Biomixing of the oceans? *Science* **316**, 838–839 (2007).
5. Sato, M., Dower, J. F., Kunze, E. & Dewey, R. Second-order seasonal variability in diel vertical migration timing of euphausiids in a coastal inlet. *Mar. Ecol. Prog. Ser.* **480**, 39–56 (2013).
6. Hamner, W. M., Hamner, P. P., Strand, S. W. & Gilmer, R. W. Behavior of antarctic krill, *Euphausia superba*: chemoreception, feeding, schooling, and molting. *Science* **220**, 433–435 (1983).
7. Bianchi, D., Galbraith, E. D., Carozza, D. A., Mislan, K. A. S. & Stock, C. A. Intensification of open-ocean oxygen depletion by vertically migrating animals. *Nat. Geosci.* **6**, 545–548 (2013).
8. Pedley, T. J. & Kessler, J. O. Hydrodynamic phenomena in suspensions of swimming microorganisms. *Annu. Rev. Fluid Mech.* **24**, 313–358 (1992).
9. Koch, D. L. & Subramanian, G. Collective hydrodynamics of swimming microorganisms: living fluids. *Annu. Rev. Fluid Mech.* **43**, 637–659 (2011).
10. Zhang, X. & Dam, H. G. Downward export of carbon by diel migrant mesozooplankton in the central equatorial Pacific. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **44**, 2191–2202 (1997).
11. Longhurst, A. R. & Glen Harrison, W. Vertical nitrogen flux from the oceanic photic zone by diel migrant zooplankton and nekton. *Deep Sea Res. A* **35**, 881–889 (1988).
12. Wilhelmus, M. M. & Dabiri, J. O. Observations of large-scale fluid transport by laser-guided plankton aggregations. *Phys. Fluids* **26**, 101302 (2014).
13. Hanamura, Y., Endo, Y. & Taniguchi, A. Underwater observations of the surface swarm of a euphausiid, *Euphasia pacifica*, in Sendai Bay, northeasterern Japan. *Mer (Paris)* **22**, 63–68 (1984).
14. Kunze, E., Dower, J. F., Beveridge, I., Dewey, R. & Bartlett, K. P. Observations of biologically generated turbulence in a coastal inlet. *Science* **313**, 1768–1770 (2006).
15. Omori, M. & Hamner, W. M. Patchy distribution of zooplankton: behavior, population assessment and sampling problems. *Mar. Biol.* **72**, 193–200 (1982).
16. Catton, K. B., Webster, D. R., Kawaguchi, S. & Yen, J. The hydrodynamic disturbances of two species of krill: implications for aggregation structure. *J. Exp. Biol.* **214**, 1845–1856 (2011).
17. Tarling, G. A. & Thorpe, S. E. Oceanic swarms of antarctic krill perform satiation sinking. *Proc. R. Soc. B* **284**, 20172015 (2017).
18. Jackson, P. R. & Rehmann, C. R. Experiments on differential scalar mixing in turbulence in a sheared, stratified flow. *J. Phys. Oceanogr.* **44**, 2661–2680 (2014).
19. Turner, J. S. Jets and plumes with negative or reversing buoyancy. *J. Fluid Mech.* **26**, 779–792 (2006).
20. Shih, L. H., Koseff, J. R., Ivey, G. N. & Ferziger, J. H. Parameterization of turbulent fluxes and scales using homogeneous sheared stably stratified turbulence simulations. *J. Fluid Mech.* **525**, 193–214 (2005).
21. Nash, J. D. & Moum, J. N. Microstructure estimates of turbulent salinity flux and the dissipation spectrum of salinity. *J. Phys. Oceanogr.* **32**, 2312–2333 (2002).
22. Aref, H. & Jones, S. W. Enhanced separation of diffusing particles by chaotic advection. *Phys. Fluids A* **1**, 470–474 (1989).
23. Moore, C. M. et al. Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**, 701–710 (2013).
24. DeLong, E. F. Microbial community genomics in the ocean. *Nat. Rev. Microbiol.* **3**, 459–469 (2005).
25. Fuhrman, J. A., Cram, J. A. & Needham, D. M. Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* **13**, 133–146 (2015).

Author contributions I.A.H., J.O.D., J.R.K. and S.G.M. conceived the irreversible mixing experiments; I.A.H. and J.O.D. conceived the flow visualization experiments; I.A.H. conducted all experiments; all authors contributed to data analysis; I.A.H. and J.O.D. wrote the initial draft of the manuscript, and all authors contributed to manuscript revisions.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Laboratory migration systems. The following systems were used to create the laboratory migrations.

LED array animal guidance. Animal guidance in the 1.2-m-tall tank consisted of two arrays of LEDs with focusing optics with approximately 250-mm focal length (Outlite a100) and blue and green filters arranged to illuminate a ten-centimetre-square area. One LED array was used at a time to induce unidirectional swimming towards the top (blue LEDs) or bottom (green LEDs) of the tank.

Laser animal guidance. Animal guidance in the 2-m-tall tank consisted of two lasers. A 1-W, 447-nm (blue) laser was mounted horizontally above the tank and aligned with a −50-mm focal length spherical lens (Newport Optics) and a 45° mirror to redirect the beam vertically downward in the centre of the water column. A 1-W, 532-nm (green) laser was mounted horizontally beneath the tank with identical optics to redirect the beam vertically upward. The blue and green laser beams were centred and aligned throughout the height of tank. One laser was used at a time to induce unidirectional swimming towards the activated laser, that is, towards the top blue laser or bottom green laser.

Animal protocol. Adult-sized *A. salina*, commonly known as brine shrimp, were obtained before each experiment (Mariculture Tech) and allowed to acclimatize in aerated beakers for at least 3 h at a salinity similar to that used in the experiment (26‰). The animals were then introduced at the top of the tank and allowed to further acclimatize before commencing an experiment. In the 1.2-m tank, animals were introduced at a tank-averaged abundance of 46,000 ± 5,000 animals per m³ to 138,000 ± 5,000 animals per m³. Irreversible mixing measurements were insensitive to the range of animals introduced to the tank in these experiments. In the 2-m tank, animals were introduced at a tank-averaged abundance of 20,000 ± 5,000 animals per m³.

Tank stratification and density measurements. All experiments used salt to produce the density stratification and held the temperature at a constant 21 °C, matching ambient laboratory conditions. For measurements of irreversible mixing, the 1.2-m-tall tank was stratified with variable-salinity water to form two layers of distinct density, similar to setups for a variety of laboratory mixing experiments in the literature[26,27]. The lighter layer (26.0‰) was first pumped into the tank followed by the denser layer (26.3 ± 0.2‰) pumped in from below. For flow visualization, the 2-m-tall tank was linearly stratified using the double bucket technique[28] between 25.6‰ of salt at the surface and 26.2% of salt at a depth of 2 m, yielding a buoyancy frequency of $N = 0.05\,\mathrm{s}^{-1}$. Animal swimming behaviour and the large-scale jet formation were unaffected by the stratification type used (that is, two-layer or linear).

A conductivity–temperature profiler (125MicroScale Conductivity and Temperature Instrument (MSCTI), Precision Measurement Engineering) was used to measure the local density with an accuracy of $\pm 0.005\,\mathrm{kg\,m}^{-3}$. To minimize probe drift due to exposure to air, the 6-cm span of the probe tip was continuously immersed in the tank, restricting measurements to 6 cm below the surface. Density profiles were linearly extrapolated 6 cm to the surface in post-processing for evaluation of mass conservation. The raw data points were used to calculate the water column mass while a spline fit with uncertainty bounds was used for the displayed density profiles. To obtain the normalized density profiles presented, the location of peak stratification of the initial density profile was found and the density and height values were shifted such that this location passed through the origin. Density was then rescaled to span [−1,1] following the expression,

$$\rho_{\mathrm{normalized}} = \frac{\rho_{\mathrm{raw}} - \rho_{\mathrm{peak}}}{(\rho_{\mathrm{max}} - \rho_{\mathrm{min}})/2} \qquad (1)$$

in which ρ_{raw} is the raw density data, ρ_{peak} is the density at the peak stratification, ρ_{max} is the highest density (density of the bottom layer) and ρ_{min} is the lowest density (density of the top layer). All subsequent profiles in the same experiment were shifted and rescaled with values (ρ_{peak}, ρ_{max} and ρ_{min}) from the initial profile.

Density profiles were obtained at the end of each ten minute downward migration cycle. For quantitative analysis of mass conservation and irreversible mixing, repeated final profiles were obtained every ten minutes after the 120-min experiment until it was confirmed that density profiles were constant in time, signalling that the tank was fully settled and horizontally homogeneous. The 120-min experiment was conducted six times. The density differences between the top and bottom layers in each experiment were $0.03\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.04\,\mathrm{s}^{-1}$), $0.05\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.05\,\mathrm{s}^{-1}$), $0.16\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.09\,\mathrm{s}^{-1}$), $0.21\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.10\,\mathrm{s}^{-1}$), $0.22\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.10\,\mathrm{s}^{-1}$) and $0.36\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.13\,\mathrm{s}^{-1}$). For the experiment with a density difference of $0.16\,\mathrm{kg\,m}^{-3}$ ($N_{\mathrm{int}} = 0.09\,\mathrm{s}^{-1}$), two 30-min upward migrations were induced instead of six 10-min upward migrations, to study the

effect of migration duration. The final density profile was consistent for the two experiments with the same cumulative upward migration time. For experiments with similar stratification strength ($N_{\mathrm{int}} = 0.09$, $N_{\mathrm{int}} = 0.10$ and $N_{\mathrm{int}} = 0.10$) the normalized density profiles collapsed onto the same shape. When overlaying the normalized density profiles, the standard deviation of the normalized density measurements for the three different experiments was approximately 6.8 per cent for the initial profiles and 6.9 per cent for the final profiles, indicating the profiles evolved similarly in each experiment.

Calculation of effective diffusivity. Depth-independent effective diffusivity, calculated from an error function fit to the final density profiles, was of order $10^{-2}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, three orders of magnitude larger than the molecular diffusivity of salt. However, a constant effective diffusivity with an initial error-function density profile yields symmetric error-function final profiles. Thus, these estimates produced poor fits to the spatial variability of the experimental data. To improve the estimate, a depth-dependent effective diffusivity was used. A numerical integrator (MATLAB, Mathworks) solved the diffusion equation,

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial z}\left(\kappa_{\mathrm{eff}}(z)\frac{\partial C}{\partial z}\right)$$

using a forward-in-time, centred-in-space explicit numerical scheme to obtain final density profiles from the initial experimental profiles and a prescribed effective diffusivity, $\kappa_{\mathrm{eff}}(z)$. A simulated annealing algorithm (MATLAB, Mathworks) was used to minimize the norm of the difference between the experimental final profile and the numerically calculated final profile with a depth-dependent effective diffusivity. This method informed the general shape of effective diffusivity with height, yielding approximately an error-function-shape effective diffusivity profile with the highest values above the pycnocline. We then empirically optimized the general shape obtained from the simulated annealing algorithm with a parameter sweep of constants defining the shift, spread and amplitude of an error function. This optimization was done to obtain a conservative estimate of effective diffusivity, that is, chosen to underestimate effective diffusion, particularly in regions where the full asymmetry of the profile was difficult to capture numerically. Peak values of the depth-dependent effective diffusivity were the same order of magnitude as the calculated depth-independent effective diffusivity.

Flow visualization. The following techniques were used for flow visualization.

Schlieren imaging. Density gradients in the flow were visualized with a schlieren imaging setup with a field of view of 5 cm in the 2-m-tall tank. A 75-W Xenon arc lamp was used to illuminate the field. The light passed through a condenser lens and horizontal slit and then was collimated via a 250-mm focal length plano-convex doublet lens (Newport Optics). The resultant 5-cm diameter parallel light beam travelled through the width of the tank at a height 1 m above the bottom of the tank and was distorted by the varying indices of refraction of the perturbed density field in the path of the light. The beam then passed through a 400-mm focal length plano-convex doublet lens that converged the beam to a focal point. A precision aligner was used to position a vertical razor blade at the focal point of the beam. For the varying density flow, the deflected rays do not coincide with the focal point and were either blocked by the razor blade or passed through without any reduction in intensity, depending on the density gradients encountered[29]. The total light distortion over the path length yielded an image with regions of higher and lower illumination corresponding to regions of the highest positive and negative fluid density gradients, respectively, in the direction normal to the razor blade edge (that is, horizontally).

Single animal schlieren (Fig. 4a and Supplementary Video 2) was conducted with the assistance and facilities of R. Strickler (University of Wisconsin, Milwaukee) with a free swimming animal in a 10 × 5 × 15-cm tank linearly stratified from 0‰ at the surface to 35‰ at a depth of 15 cm. The schlieren setup was similar to that described above except with 250-mm and 150-mm focal length plano-convex doublet lenses creating a 1.5-cm field of view. The considerably stronger density stratification and smaller field of view produced a much stronger and clearer schlieren signal than in the larger-scale experiments with no observable effect on swimming behaviour.

Tracer particle overlay. The 2-m-tall tank was seeded with 10-μm glass beads and illuminated with a vertical red laser sheet formed with a cylindrical lens with a focal length $f = -9.7\,\mathrm{mm}$ (Newport Optics) resulting in a vertical sheet height of 40 cm at the location of the migration. *Artemia* response to red wavelengths is minimal, so animal behaviour was unaffected by the additional laser. A camera (Nikon D750, 24 megapixel) with a focal length $f = 14$-mm lens (Nikon AF Nikkor) was used to record video focused to the right of the migration when viewed on camera. The wide field of view of the lens resulted in image distortion that was removed during post-processing. Sets of 250 consecutive frames, corresponding to 4.2 s of video, were overlaid by storing the maximum illumination of each pixel in the stack resulting in streaks representing the pathlines of particles in the flow. Swimming animals in the frames resulted in wider grey streaks much larger than the actual

individual size. Animals from the final frame were overlaid and highlighted in red to illustrate relative size.

Planar-laser-induced fluorescence. Planar-laser-induced fluorescence was used to obtain an aggregation-scale view of the resultant hydrodynamics in the 1.2-m tank. A 532-nm (green) laser with an $f = -25$-mm plano-concave cylindrical lens (Newport Optics) created a laser sheet illuminating fluorescent dye (Rhodamine 6 G, Sigma-Aldrich) in the tank. Animals had a stronger phototactic response to the blue LED array, so introduction of the green sheet laser during an upward migration did not affect animal behaviour. Qualitative visualization of the transport through an animal aggregation was conducted with 3 ml of 500 p.p.m. Rhodamine 6 G dye slowly introduced at the water surface with a horizontally oriented pipette such that the dye had no initial vertical momentum. In a quiescent tank (or with the aggregation at the bottom of the tank), the buoyant dye spread into a very thin layer at the top of the tank and remained there with very minor diffusion ($D_{Rh6G} \approx 10^{-10}\,\mathrm{m^2\,s^{-1}}$). As an upward migration developed, the animals propelled fluid rearward, eventually transporting the dye from the surface downward through the entire extent of the migration. An orange filter placed over the camera (Nikon D750, 24 megapixel) isolated the fluorescent signal from the dye. Flow patterns were recorded in real-time at 60 frames per second as experiments were carried out.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. Density profile data are available at https://github.com/ihoughton/Houghton2018_data. All other data are available from the corresponding author upon reasonable request.

26. Linden, P. F. Mixing across a density interface produced by grid turbulence. *J. Fluid Mech.* **100**, 691–703 (2006).
27. Whitehead, J. A. & Stevenson, I. Turbulent mixing of two-layer stratified fluid. *Phys. Fluids* **19**, 125104 (2007).
28. Economidou, M. & Hunt, G. R. Density stratified environments: the double-tank method. *Exp. Fluids* **46**, 453–466 (2009).
29. Settles, G. S. *Schlieren and Shadowgraph Techniques* (Springer, Berlin, 2001).

LETTER

Electrophilic properties of itaconate and derivatives regulate the IκBζ–ATF3 inflammatory axis

Monika Bambouskova[1], Laurent Gorvel[1], Vicky Lampropoulou[1], Alexey Sergushichev[2], Ekaterina Loginicheva[1], Kendall Johnson[3], Daniel Korenfeld[1], Mary Elizabeth Mathyer[4], Hyeryun Kim[3], Li-Hao Huang[1], Dustin Duncan[5], Howard Bregman[3], Abdurrahman Keskin[6], Andrea Santeford[7], Rajendra S. Apte[7], Raghav Sehgal[8], Britney Johnson[1], Gaya K. Amarasinghe[1], Miguel P. Soares[9], Takashi Satoh[10], Shizuo Akira[10], Tsonwin Hai[11], Cristina de Guzman Strong[4], Karine Auclair[5], Thomas P. Roddy[3], Scott A. Biller[3], Marko Jovanovic[6], Eynav Klechevsky[1], Kelly M. Stewart[3], Gwendalyn J. Randolph[1] & Maxim N. Artyomov[1]*

Metabolic regulation has been recognized as a powerful principle guiding immune responses. Inflammatory macrophages undergo extensive metabolic rewiring[1] marked by the production of substantial amounts of itaconate, which has recently been described as an immunoregulatory metabolite[2]. Itaconate and its membrane-permeable derivative dimethyl itaconate (DI) selectively inhibit a subset of cytokines[2], including IL-6 and IL-12 but not TNF. The major effects of itaconate on cellular metabolism during macrophage activation have been attributed to the inhibition of succinate dehydrogenase[2,3], yet this inhibition alone is not sufficient to account for the pronounced immunoregulatory effects observed in the case of DI. Furthermore, the regulatory pathway responsible for such selective effects of itaconate and DI on the inflammatory program has not been defined. Here we show that itaconate and DI induce electrophilic stress, react with glutathione and subsequently induce both Nrf2 (also known as NFE2L2)-dependent and -independent responses. We find that electrophilic stress can selectively regulate secondary, but not primary, transcriptional responses to toll-like receptor stimulation via inhibition of IκBζ protein induction. The regulation of IκBζ is independent of Nrf2, and we identify ATF3 as its key mediator. The inhibitory effect is conserved across species and cell types, and the in vivo administration of DI can ameliorate IL-17–IκBζ-driven skin pathology in a mouse model of psoriasis, highlighting the therapeutic potential of this regulatory pathway. Our results demonstrate that targeting the DI–IκBζ regulatory axis could be an important new strategy for the treatment of IL-17–IκBζ-mediated autoimmune diseases.

Differential gene expression in our published transcriptional analysis of DI-treated bone marrow-derived macrophages (BMDMs)[2] showed the enrichment of electrophilic and xenobiotic stress response pathways, specifically the upregulation of classical transcriptional markers of a Nrf2-mediated response[4] such as *Hmox1*, *Nqo1* and *Gclm* (Fig. 1a). Indeed, Nrf2 protein levels, as well as the protein levels of its classical target genes, increased during 12 hours of treatment with DI (Fig. 1b). Endogenous itaconate also induced a KEAP1–Nrf2 response, as Nrf2 protein levels upon lipopolysaccharide (LPS) activation were higher in wild-type macrophages than in *Irg1*$^{-/-}$ macrophages (*Irg1* is also known as *Acod1*) (Fig. 1c). Overall, transcriptional signatures of DI treatment matched those of the KEAP1-knockout macrophages[5] (Extended Data Fig. 1a), indicating the induction of a global electrophilic stress response.

Considering its structure, DI can readily act as an electrophile in Michael reactions (Extended Data Fig. 1b) and trigger electrophilic stress, which is usually controlled via glutathione (GSH) buffering. For example, covalent conjugation with GSH has been described for fumarate, a metabolite that typically accumulates in cells with mutated fumarate hydratase[6,7]. Analysis of the cell media from DI-treated macrophages suggested that DI was taken up from the media, and a substantial peak was found with a retention time of 14.4 min and an *m/z* of 464.1334; this peak increased during incubation with DI and corresponded to the diester of methylsuccinated GSH (DI-GSH) (Fig. 1d, e, Extended Data Fig. 1c, d). The origin of DI-GSH was confirmed using synthesized $^{13}C_5$-labelled DI for cell treatment (Extended Data Fig. 1e). Notably, the reactivity with GSH was also observed for natural itaconate; we detected methylsuccinated GSH (Ita-GSH, Fig. 1d) in LPS-stimulated macrophages, which correlated with itaconate production and was absent in *Irg1*$^{-/-}$ cells (Fig. 1f). The identities of the DI-GSH and Ita-GSH conjugates were confirmed by comparing their liquid chromatography–mass spectrometry retention times with those of chemically synthesized standards (Extended Data Fig. 1f, g).

The reactivity of DI led to a substantial decrease in cellular GSH concentration and an associated increase in the generation of reactive oxygen species (Fig. 1g, Extended Data Fig. 1h). Therefore, we tested a panel of antioxidants (scavengers of reactive oxygen species) by administering them simultaneously with DI. Co-treatment with only *N*-acetylcysteine (NAC) or cell-permeable GSH (EtGSH) reversed the effect of DI, suggesting that DI acts preferentially by modulating the cellular pool of thiol-containing molecules (Fig. 1h, Extended Data Fig. 1i). The prototypical electrophile dimethyl fumarate (DMF) also triggered an Nrf2 response and showed selective inhibition of IL-6 (Fig. 1i, Extended Data Fig. 1j).

Tnf is induced during the primary transcriptional response to toll-like receptor (TLR) stimulation, whereas *Il6* is a product of the secondary transcriptional responses[8]. The major transcription factor that has been reported to selectively regulate secondary transcriptional response to TLR activation is IκBζ, which is encoded by the *Nfkbiz* gene[9] (Fig. 2a, Extended Data Fig. 2a, b). Notably, DI completely abolished the LPS-associated induction of IκBζ protein and inhibited its target genes (including *Il12b* and *Edn1*) in both BMDMs and human blood monocytes (Fig. 2b, Extended Data Fig. 2c–g). This suggested that the observed specificity of action of DI on IL-6 may result from the selective inhibition of the secondary transcriptional response to TLR activation. Indeed, DI did not inhibit IκBα degradation and upstream signalling in response to LPS, nor did it prevent LPS-mediated p65 nuclear translocation (Extended Data Fig. 2h–j). We next tested whether the NAC- or EtGSH-mediated

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA. [2]Computer Technologies Department, ITMO University, Saint Petersburg, Russia. [3]Agios Pharmaceuticals, Cambridge, MA, USA. [4]Division of Dermatology, Center for Pharmacogenomics, Center for the Study of Itch, Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA. [5]Department of Chemistry, McGill University, Montreal, Quebec, Canada. [6]Department of Biological Sciences, Columbia University, New York, NY, USA. [7]Department of Ophthalmology and Visual Sciences, Washington University School of Medicine, St. Louis, MO, USA. [8]Elucidata Corporation, Cambridge, MA, USA. [9]Instituto Gulbenkian de Ciência, Oeiras, Portugal. [10]Host Defense, Immunology Frontier Research Center, Osaka University, Suita, Japan. [11]Department of Biological Chemistry and Pharmacology, Ohio State University, Columbus, OH, USA. *e-mail: martyomov@wustl.edu
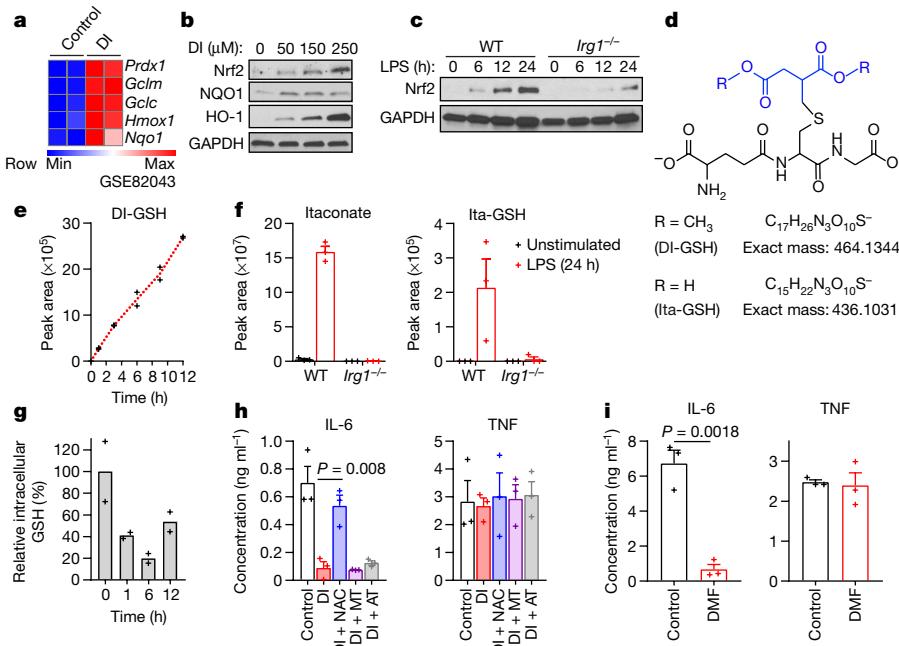
Fig. 1 | DI and itaconate induce electrophilic stress in macrophages. a, Nrf2-response genes in DI-treated BMDMs. **b, c**, Western blots of Nrf2 and Nrf2 targets in DI-treated (**b**) or LPS-stimulated (**c**) BMDMs. **d**, The chemical structures of DI-GSH and Ita-GSH. **e**, DI-GSH levels in the media of DI-treated BMDMs, mean of $n = 2$ cultures. **f**, Itaconate (left) and Ita-GSH (right) levels in BMDMs, mean ± s.e.m., $n = 3$ cultures. **g**, GSH levels in DI-treated BMDMs, mean of $n = 2$ experiments. **h, i**, Cytokine levels in BMDMs treated as indicated and LPS-stimulated for 4 h (**h**) and 24 h (**i**); mean ± s.e.m., $n = 3$ experiments. Western blots are representatives of three experiments. For gel source data, see Supplementary Fig. 1. Statistical tests used were two-tailed t-tests. AT, α-tocopherol; MT, MitoTEMPO.

reversal of the effect of DI on IL-6 was associated with the recovery of IκBζ protein induction. Indeed, co-treatment of BMDMs with DI and NAC (or EtGSH), but not α-tocopherol, restored LPS-mediated IκBζ induction in BMDMs and in human blood monocytes, whereas NAC itself did not increase IL-6 levels in untreated or DI-treated *Nfkbiz⁻/⁻* BMDMs (Fig. 2c, Extended Data Fig. 2k–m). Consistent with the idea that the primary NF-κB-mediated response is not affected by DI at this concentration range, *Nfkbiz* mRNA levels were unchanged upon DI treatment, meaning that IκBζ protein induction is affected at the post-transcriptional level (Fig. 2d, e, Extended Data Fig. 3a).

To understand the regulation of IκBζ protein, we used proteasome inhibitor MG132 and autophagosome–lysosome fusion inhibitor bafilomycin A, but neither rescued IκBζ induction during DI treatment (Extended Data Fig. 3b). Additionally, the action of DI could not be attributed to the regulation of *Nfkbiz* 3′ untranslated region (UTR)[10], as indicated by a GFP reporter assay (Extended Data Fig. 3c). Finally, cellular stress has been shown to regulate mRNA translation via phosphorylation-driven inactivation of eIF2α[11]. Indeed, we detected a marked increase in macrophage eIF2α phosphorylation in response to DI (Extended Data Fig. 3d), suggesting that the inhibitory effect of DI on IκBζ protein levels is associated with the regulation of IκBζ
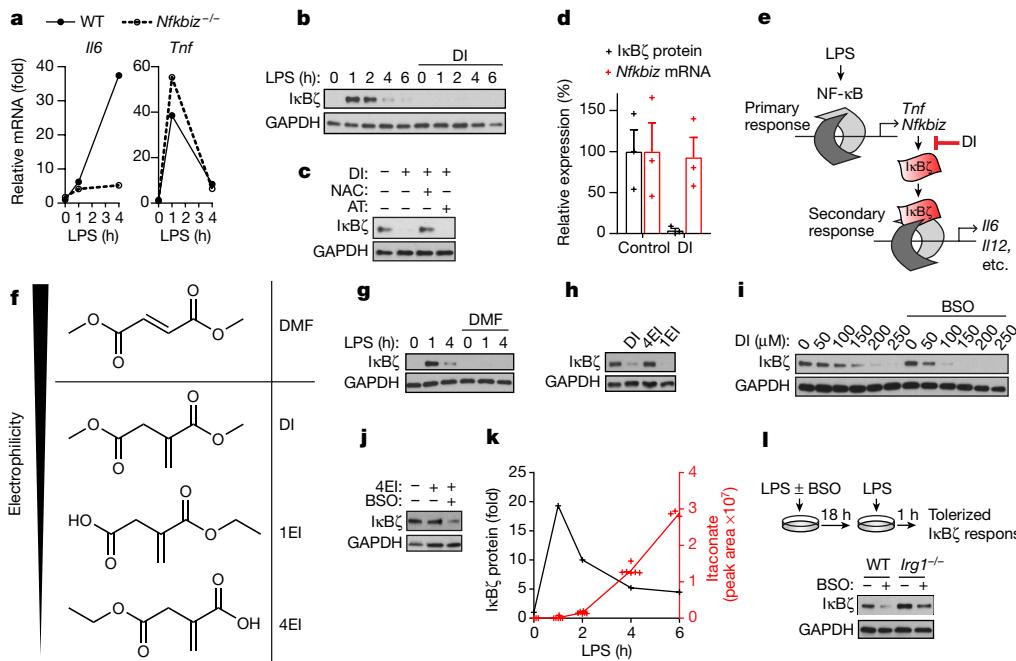
Fig. 2 | DI inhibits LPS-mediated IκBζ induction. a, mRNA expression in LPS-stimulated BMDMs. Representative of two experiments. **b, c, g–j**, Western blots of IκBζ expression in BMDMs treated with DI, DMF, 4EI (**h**, 5 μM; **j**, 10 μM) or 1EI, LPS for 1 h or as indicated. **d**, Relative levels of IκBζ protein and mRNA in BMDMs treated with DI and then LPS for 1 h, mean ± s.e.m., $n = 3$ experiments. **e**, Schematic showing the mechanism of action of DI. **f**, Structures of DMF and itaconate derivatives. **k**, Densitometry of IκBζ from **b** and itaconate levels, mean of $n = 6$ cultures. **l**, Schematic (top) and western blot (bottom) for the measurement of IκBζ expression in BMDMs tolerized in the presence of BSO. Western blots are representative of three experiments (except for **h**, two experiments). For gel source data, see Supplementary Fig. 1.
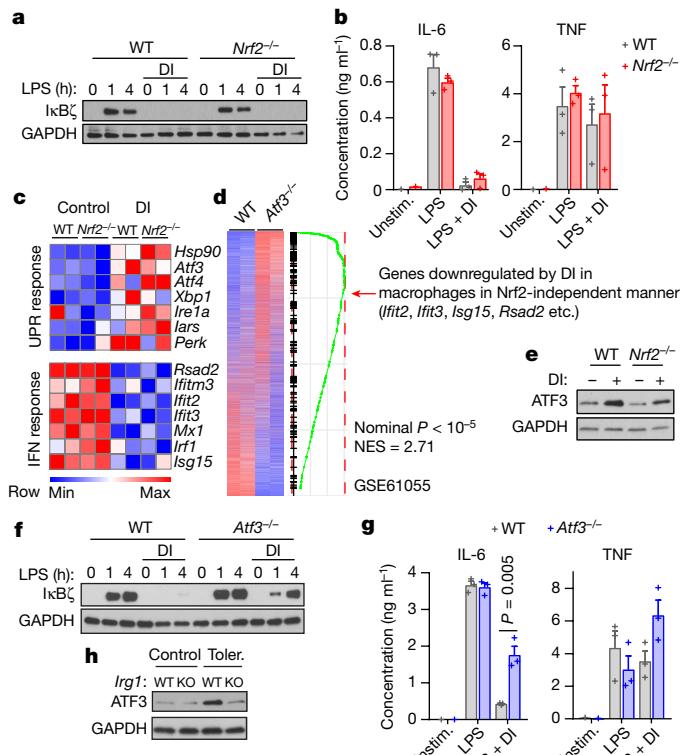
Fig. 3 | DI induces an Nrf2-independent response and inhibits the IL-6–IκBζ axis via ATF3. a, f, Western blots showing IκBζ expression in *Nrf2*[−/−] (**a**) and *Atf3*[−/−] (**f**) BMDMs. **b, g,** Cytokine levels in BMDMs treated with DI, stimulated with LPS for 4 h (**b**) and 24 h (**g**), mean ± s.e.m., *n* = 3 experiments. **c,** Genes regulated by DI independently of Nrf2. *Hsp90*, *Ire1a* and *Perk* are also known as *Hsp84-2*, *Ern1* and *Eif2ak3*, respectively. 'UPS response' indicates UPS response pathways. **d,** Transcriptional comparison of *Atf3*[−/−] and wild-type BMDMs and enrichment of the Nrf2-independent DI signature. **e, h,** Western blots of ATF3 expression in BMDMs after DI treatment (**e**) and tolerized in the presence of BSO (**h**). Western blot data are representative of three experiments. For gel source data, see Supplementary Fig. 1. Statistical tests used were two-tailed *t*-tests. KO, knockout; NES, normalized enrichment score; UPR, unfolded protein response.

translation. However, this suppression was very specific to IκBζ and possibly a few other proteins; DI did not affect protein synthesis globally, as judged by either metabolic labelling of nascent proteins or proteomic profiling of DI-treated macrophages (Extended Data Fig. 3e–h).

To further understand the effect of electrophilic stress on the IκBζ-mediated inflammatory program, we tested a panel of itaconate derivatives with different electrophilicities and also included DMF (Fig. 2f). The inhibitory effects of the compounds on IκBζ correlated with their electrophilic strength: DI, DMF and 1-ethyl itaconate (1EI) inhibited IκBζ induction, whereas 4-ethyl itaconate (4EI) did not (Fig. 2g, h). To decouple the contribution of the feedback GSH synthesis during the electrophilic stress response to these compounds, we tested combinations of itaconate derivatives with buthionine sulfoximine (BSO), a non-electrophilic inhibitor of GSH synthesis, which itself neither inhibits cytokine production nor triggers a Nrf2 response (Extended Data Fig. 4a–c). BSO enhanced the inhibitory effect of DI on IκBζ–IL-6 (Fig. 2i, Extended Data Fig. 4d) and unlocked the effect of 4EI on IκBζ–IL-6 (Fig. 2j, Extended Data Fig. 4e).

We next considered the effects of endogenously produced itaconate, which is a weak electrophile, on the IκBζ-regulatory axis. The temporal dynamics of IκBζ upon LPS activation are very different from those of itaconate: levels of IκBζ reach a maximum at around 1 hour and are already reduced considerably by around 4 hours, whereas itaconate is only induced at around 2–4 hours and plateaus later after LPS
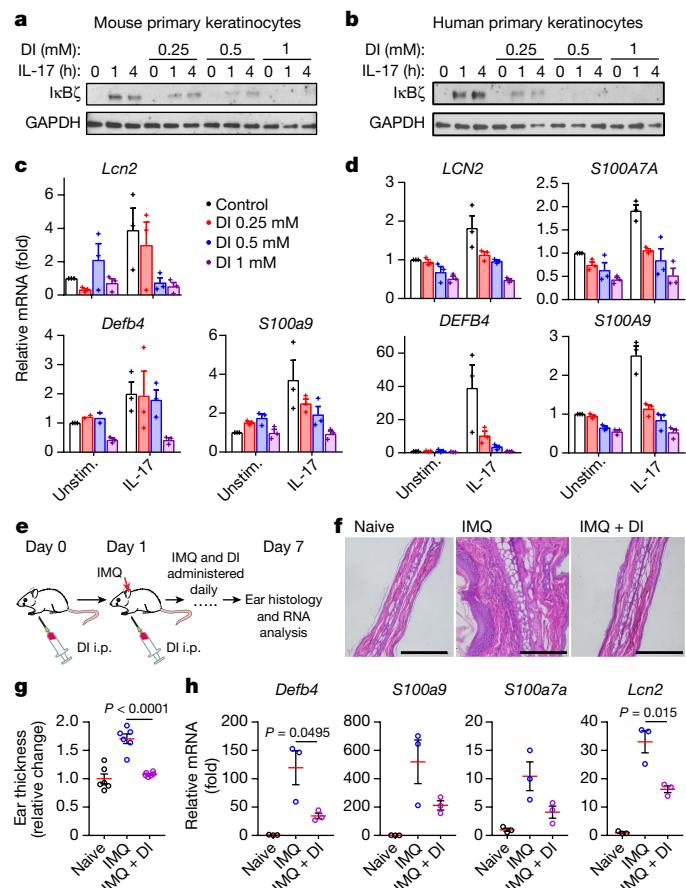
Fig. 4 | DI inhibits IL-17-mediated IκBζ induction in keratinocytes and ameliorates psoriatic pathology. a, b, Western blot of IκBζ expression in DI-treated, IL-17A-stimulated primary mouse (**a**) and human (**b**) keratinocytes. Representative of three mice or donors. For gel source data, see Supplementary Fig. 1. **c, d,** mRNA expression in DI-treated, IL-17A-stimulated (4 h) primary mouse (**c**) and human (**d**) keratinocytes, mean ± s.e.m., *n* = 3 mice or donors. **e,** Schematic of DI administration in a psoriasis model. **f,** Ear histology of control mice (left), IMQ-treated mice (middle) and mice treated with both IMQ and DI (right). Scale bars, 100 μm. Representative of six mice in two experiments. **g,** Quantification of the histology results in **f**, showing the relative change in ear thickness. Mean ± s.e.m., *n* = 6 mice. **h,** mRNA expression to show the induction of the indicated IκBζ target genes in ear tissue, mean ± s.e.m., *n* = 3 mice. Statistical tests used were two-tailed *t*-tests.

stimulation (Fig. 2k). As such, to address the physiological relevance of endogenous itaconate to IκBζ regulation, we designed an experiment in which both natural itaconate and IκBζ are present in the cells at the same time. Specifically, we have evaluated in vitro macrophage tolerization, when cells are first stimulated with LPS and then rechallenged with LPS at 18 hours. This design ensures that endogenous itaconate is produced in sufficient amounts and LPS restimulation triggers IκBζ induction once again. We observed that, in the presence of BSO, there was a marked difference in IκBζ protein levels upon restimulation between *Irg1*[−/−] and wild-type cells (Fig. 2l, Extended Data Fig. 4f), confirming the potential regulatory effects of natural itaconate on IκBζ.

We next aimed to identify major regulatory hubs connecting the electrophilic stress to the blockade of IκBζ induction. First, we tested whether Nrf2 is involved in the inhibition of IκBζ synthesis by DI. Nrf2-deficient BMDMs did not alleviate DI-mediated IκBζ and IL-6 inhibition (Fig. 3a, b, Extended Data Fig. 5a). Similarly, macrophages genetically lacking p62 (also known as SQSTM1)[12] or HO-1[13] did not rescue DI-mediated IκBζ inhibition (Extended Data Fig. 5b–d). Because the action of DI is independent of Nrf2, we performed RNA sequencing (RNA-seq) in *Nrf2*[−/−] (also known as *Nfe2l2*[−/−]) and wild-type BMDMs and analysed the genes that were differentially

expressed upon treatment with DI (Extended Data Fig. 5e). Most of the differentially DI-regulated pathways of the integrated stress response were upregulated (*Atf3*, *Atf4* and *Eif2ak3* (also known as *Perk*)), whereas the interferon (IFN) response pathways (for example, *Isg15*) were consistently downregulated (Fig. 3c, Extended Data Fig. 5f). Focusing on ATF3 as a potential candidate[14], we compared our Nrf2-independent transcriptional signature of DI treatment to the publicly available dataset that profiled wild-type and *Atf3*[−/−] macrophages at their basal states[15]. Notably, we found highly statistically significant overlap between genes regulated by ATF3 and genes regulated by DI (Fig. 3d), suggesting that the action of DI might be mediated by ATF3. Indeed, the protein ATF3 was upregulated by DI treatment even on an *Nrf2*[−/−] background (Fig. 3e). Furthermore, *Atf3*[−/−] cells restored IκBζ protein levels upon DI treatment and significantly increased IL-6 production in DI-treated *Atf3*[−/−] cells compared to the wild type (Fig. 3f, g). In *Atf3*[−/−] macrophages, DI failed to increase eIF2α phosphorylation, even though it still induced an Nrf2 response (Extended Data Fig. 5g, h). DI-mediated ATF3 expression was efficiently decreased by co-treatment with NAC or EtGSH in both mouse macrophages and human monocytes (Extended Data Fig. 5i–k). Endogenous itaconate also efficiently induced an ATF3 response: when we compared wild-type and *Irg1*[−/−] BMDMs tolerized in the presence of BSO and then restimulated, ATF3 was induced in wild-type cells but not in *Irg1*[−/−] cells (Fig. 3h).

Notably, IκBζ also has a major role outside the macrophage context: it is induced upon IL-17 treatment of epithelial cells and orchestrates downstream inflammatory responses[16–18]. *Nfkbiz* polymorphisms have been associated with several immune-related conditions, including psoriasis[19]. As such, we first tested the in vitro effect of DI pretreatment on IκBζ induction in IL-17A-stimulated primary keratinocytes. The induction of IκBζ was inhibited by DI in primary mouse and human keratinocytes (Fig. 4a, b, Extended Data Fig. 6). To further examine the inhibitory effect of DI on IκBζ induction by IL-17, we analysed expression of the well-characterized IκBζ target genes *Defb4*, *S100a7a*, *Lcn2* and *S100a9* in mouse and human keratinocytes. As expected, the expression of these genes was downregulated by DI in correlation with IκBζ protein levels (Fig. 4c, d). These data suggest that DI can also modulate the induction of IκBζ in several immune contexts.

We next explored the ability of DI to interfere with IκBζ signalling in vivo. We used a mouse model of psoriasis induced by the TLR7/8 agonist imiquimod (IMQ). In this model, skin inflammation is induced by topical application of IMQ cream onto the ears of mice to induce a psoriasis-like pathology resembling the human disease. Daily topical application of IMQ to the skin of the ear for seven days led to considerable scaling and oedema of the skin in control animals, whereas mice treated with DI in addition to IMQ showed no detectable skin changes (Fig. 4e–g). Similarly, quantitative analysis of ear-skin-derived mRNA showed significant induction of the IκBζ target genes *Defb4, S100a9, S100a7a* and *Lcn2* after IMQ application, whereas their expression was markedly reduced in the skin of mice treated with DI (Fig. 4h). Daily DI administration did not affect succinate dehydrogenase activity in the heart and the liver considerably (Extended Data Fig. 7), suggesting a favourable safety profile. These data demonstrate that DI can act as an IκBζ inhibitor in vivo and could provide a targeted approach for treating various autoimmune conditions.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0052-z.

1. Jha, A. K. et al. Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization. *Immunity* **42**, 419–430 (2015).
2. Lampropoulou, V. et al. Itaconate links inhibition of succinate dehydrogenase with macrophage metabolic remodeling and regulation of inflammation. *Cell Metab.* **24**, 158–166 (2016).
3. Cordes, T. et al. Immunoresponsive gene 1 and itaconate inhibit succinate dehydrogenase to modulate intracellular succinate levels. *J. Biol. Chem.* **291**, 14274–14284 (2016).
4. Gorrini, C., Harris, I. S. & Mak, T. W. Modulation of oxidative stress as an anticancer strategy. *Nat. Rev. Drug Discov.* **12**, 931–947 (2013).
5. Kobayashi, E. H. et al. Nrf2 suppresses macrophage inflammatory response by blocking proinflammatory cytokine transcription. *Nat. Commun.* **7**, 11624 (2016).
6. Sullivan, L. B. et al. The proto-oncometabolite fumarate binds glutathione to amplify ROS-dependent signaling. *Mol. Cell* **51**, 236–248 (2013).
7. Zheng, L. et al. Fumarate induces redox-dependent senescence by modifying glutathione metabolism. *Nat. Commun.* **6**, 6001 (2015).
8. Medzhitov, R. & Horng, T. Transcriptional control of the inflammatory response. *Nat. Rev. Immunol.* **9**, 692–703 (2009).
9. Yamamoto, M. et al. Regulation of Toll/IL-1-receptor-mediated gene expression by the inducible nuclear protein IκBζ. *Nature* **430**, 218–222 (2004).
10. Dhamija, S. et al. IL-1-induced post-transcriptional mechanisms target overlapping translational silencing and destabilizing elements in IκBζ mRNA. *J. Biol. Chem.* **285**, 29165–29178 (2010); erratum **291**, 24801 (2016).
11. Harding, H. P. et al. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol. Cell* **11**, 619–633 (2003).
12. Komatsu, M. et al. The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of KEAP1. *Nat. Cell Biol.* **12**, 213–223 (2010).
13. Otterbein, L. E., Soares, M. P., Yamashita, K. & Bach, F. H. Heme oxygenase-1: unleashing the protective properties of heme. *Trends Immunol.* **24**, 449–455 (2003).
14. Gilchrist, M. et al. Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* **441**, 173–178 (2006).
15. Labzin, L. I. et al. ATF3 is a key regulator of macrophage IFN responses. *J. Immunol.* **195**, 4446–4455 (2015).
16. Okuma, A. et al. Enhanced apoptosis by disruption of the STAT3-IκB-ζ signaling pathway in epithelial cells induces Sjögren's syndrome-like autoimmune disease. *Immunity* **38**, 450–460 (2013).
17. Muromoto, R. et al. IL-17A plays a central role in the expression of psoriasis signature genes through the induction of IκB-ζ in keratinocytes. *Int. Immunol.* **28**, 443–452 (2016).
18. Johansen, C. et al. IκBζ is a key driver in the development of psoriasis. *Proc. Natl Acad. Sci. USA* **112**, E5825–E5833 (2015).
19. Tsoi, L. C. et al. Enhanced meta-analysis and replication studies identify five new psoriasis susceptibility loci. *Nat. Commun.* **6**, 7001 (2015).

Author contributions M.B. and M.N.A. conceived and designed the study and wrote the manuscript. M.B. performed western blot, cytokine, cytometry and glutathione analyses, metabolic labelling of protein synthesis and succinate dehydrogenase activity assays. L.G., D.K., M.B. and E.K. designed and performed human blood monocyte experiments. V.L., L.-H.H. and G.J.R. designed and performed in vivo psoriasis model experiments. A.Se. performed RNA-seq data analysis. E.L. prepared RNA-seq libraries and helped with PCR experiments. M.E.M. and C.d.G.S. designed and performed the isolation of mouse and human primary keratinocytes. H.K., K.J., H.B., T.P.R., S.A.B. and K.M.S. designed and performed mass spectrometry metabolomic measurements, and analysis and synthesis of Ita-GSH and DI-GSH conjugates. D.D. and K.A. synthesized [13]C5-labelled DI. A.K. and M.J. designed and performed proteomic analysis. A.Sa., R.S.A., T.H., M.P.S., T.S. and S.A. provided animals and bones for the study. R.S. helped with the analysis of the metabolic data. B.J. and G.K.A. designed and performed the initial Nrf2 experiments.

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Experimental animals. C57BL/6N wild-type mice were obtained from Charles River Laboratories. $Nrf2^{-/-}$ mice (cat. no. 017009) and control wild-type C57BL/6J mice (cat. no. 000664) were purchased from The Jackson Laboratory. $Irg1^{-/-}$ mice were published previously[2]. $Nfkbiz^{-/-}$ mice[9] were provided by S.A., sex-matched animals were used in experiments. p62-deficient mice[20] were provided by H. Virgin. Mice were maintained at Washington University under specific pathogen-free conditions in accordance with federal and university guidelines and protocols approved by the Animal Studies Committee of Washington University. Femurs and tibias from $Hmox1^{lox/-}$ and control $LyzM^{cre/cre}Hmox1^{lox/-}$ mice[21] were provided by M.P.S.. Femurs and tibias from $Atf3^{-/-}$ mice[22] and control C57BL/6 wild-type mice were provided by T.H.. Mice used for the study were 6–12 weeks old. Unless stated otherwise, female mice were used.

Bone marrow-derived macrophages and mouse cell cultures. BMDMs were prepared from 6- to 12-week-old mice as described[1] and cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine and 100 U ml^{-1} penicillin–streptomycin and mouse recombinant macrophage colony stimulating factor (20 ng ml^{-1}, Peprotech). For experiments, cells were seeded at a concentration of 10^6 cells per ml in tissue-culture plates of various formats. The cells were treated with DI (250 μM unless stated otherwise, cat. no. 592498, Sigma), DMF (50 μM unless stated otherwise, cat. no. 242926, Sigma), 3-(ethoxycarbonyl) but-3-enoic acid (1-ethyl itaconate, 1EI; 5 μM, Aris Pharmaceuticals Inc.), or 4-ethoxy-2-methylene-4-oxobutanoic acid (4-ethyl itaconate, 4EI; Aris Pharmaceuticals Inc.) for 12 h and activated with lipopolysaccharide (LPS; 100 ng ml^{-1}, *Escherichia coli* 0111:B4, Sigma) for 1 h or as indicated. In some experiments cells were stimulated with a combination of LPS (20 ng ml^{-1}) and IFNγ (50 ng ml^{-1}, Peprotech). In some experiments cells were treated with α-tocopherol (10 μM, Sigma), MitoTEMPO (500 μM, Sigma), N-acetylcysteine (1 mM, Sigma), the ethylester of GSH (EtGSH; 1 mM, Santa Cruz Biotechnology) or buthionene sulfoximine (BSO; 500 μM, Sigma) alone or in combination with DI for 12 h. In some experiments BMDMs were treated with DI for 12 h and bafilomycin A (100 nM, Sigma) or MG132 (10 μM, Selleckchem) were added 30 min before subsequent LPS stimulation.

In tolerization experiments, cells were stimulated with LPS (100 ng ml^{-1}) for 18 h then washed with PBS at 37 °C and restimulated with LPS (100 ng ml^{-1}) for 1 h. In some samples, BSO (500 μM) was present during the first stimulation.

The BV2 microglial cell line was a gift from H. Virgin. BV2 cells were maintained in DMEM medium supplemented with 10% FBS, 2 mM L-glutamine, 1 mM sodium pyruvate and 100 U ml^{-1} penicillin–streptomycin. BV2 responsiveness to TLR stimulation was tested and cells were not tested for mycoplasma contamination.

RNA sequencing analysis. mRNA was extracted with oligodT beads (Invitrogen), and libraries were prepared and quantified as described previously[23]. All RNA-seq experiments were performed in $n = 2$ independent cultures. Raw and processed data were deposited in the Gene Expression Omnibus. Pre-ranked gene set enrichment analysis (GSEA) was performed using the fgsea R package[24]. For the analysis of wild-type and $Nrf2^{-/-}$ BMDMs, genes were ranked according to signal-to-noise statistics; only the top 10,000 genes ordered by mean expression were considered. MSigDB C2 and H gene set collections were used. For the analysis of KEAP1 conditional knockout (KpCKO) (GSE71263)[5] and $Atf3^{-/-}$ (GSE61055)[15] datasets, differential expression analysis was carried out using the limma package[25], genes were ranked by the corresponding test statistics and P values were calculated using the GSEA method in the fgsea R package[24] with 200,000 gene-set permutations. Heat maps were generated using the Phantasus online service (https://artyomovlab.wustl.edu/phantasus/).

Western blots. Cells were lysed in RIPA lysis buffer system (Santa Cruz Biotechnology) and heat-denatured at 95 °C for 5 min in reducing sample buffer (Bio-Rad). Proteins were separated on 4–20% polyacrylamide gradient gels (Bio-Rad) and transferred onto PVDF membranes (0.45 μm pore size, Millipore). Non-specific binding was blocked with 5% skim milk (or 5% BSA when phosphoproteins were analysed), and membranes were probed with primary antibodies specific to Nrf2 (#12721), HO-1 (#70081), IκBζ (mouse-specific, #93726), IκBζ (#9244), ATF3 (#D2Y5W), IRAK1 (#4504; sensitivity of IRAK1 detection diminishes upon IRAK1 K63 ubiquitination[26]), phospho-IKK (Ser176/180, #2697), p62 (also known as SQSTM1) (#5114), phospho-eIF2α (Ser51, #9721), eIF2α (#5324) from Cell Signaling; glyceraldehyde 3-phosphate dehydrogenase (GAPDH; sc-25778), IκBα (sc-1643), ATF3 (sc-188), succinate dehydrogenase complex, subunit A (SDHA; sc-166909) from Santa Cruz Biotechnology; NQO1 (ab28947) from Abcam, followed by incubation with anti-rabbit-HRP (1:10,000; sc-2030) or anti-mouse-HRP (1:10,000; sc-2031) from Santa Cruz Biotechnology and Clarity Western ECL substrate (Bio-Rad). Membranes were exposed to X-ray

films (Research Products International) and developed using an SRX-101A film processor (Konica Minolta). GAPDH run on the same blot was used as a loading control. After the scanning of original films, the image brightness of some blots was adjusted and bands were cropped using ImageJ[27,28]. Densitometry was performed using ImageJ. For gel source data (unadjusted and uncropped images containing molecular size marker), see Supplementary Fig. 1.

DI detection with gas chromatography–mass spectrometry (GC–MS). For DI measurement, medium was collected from cells at various time points during incubation with 250 μM DI, and placed on ice. An equal volume of ethyl acetate (Sigma) was added and the samples were vortexed at 4 °C for 1 min. After centrifugation at 14,000g for 2 min at 4 °C, the organic phase was collected and approximately 20–30 mg of sodium sulfate (Sigma) was added. The samples were vortexed before analysis. Analysis was performed using a TRACE 1300 GC equipped with a 30-m DB-35MS capillary column connected to a Thermo TSQ Quantum MS operating under electron impact ionization at 70 eV. 1 μl of sample was injected in splitless mode at 270 °C, using helium as the carrier gas at a flow rate of 1 ml min^{-1}. The GC oven temperature was held at 100 °C for 3 min and increased to 240 °C at 3.5° min^{-1}. The MS source and the quadrupole were held at 230 °C and 280 °C, respectively, and the detector recorded ion abundance in the range 30–800 m/z.

Metabolite profiling with liquid chromatography–mass spectrometry (LC–MS). Bone marrow-derived macrophages were seeded in 96-well plates at 10^5 cell per well for all analyses. After treatment, media was removed from the wells and the cells were washed three times with PBS (37 °C) and immediately placed on dry ice. The frozen cells or media were stored at −80 °C until extraction. Cell extracts were prepared by adding 180 μl of 70/30 ethanol/H$_2$O solution at 70 °C, with 300 ng ml^{-1} $^{13}C_5$$^{15}N_1$ d$_5$-glutamate as the internal standard. After rigorous mixing and centrifugation (21,694g for 10 min at 4 °C), the supernatant was collected and transferred to another 96-well plate, and the solvent was evaporated under reduced pressure using EZ-2 (Genevac). Before injection, dried extracts were reconstituted, or medium was diluted (1:20) in LC–MS grade water. The extracted samples were analysed by high-resolution accurate mass lLC–MS. LC separation was achieved by reverse-phase ion-pairing chromatography. The system consisted of a Vanquish (Thermo Fisher Scientific) pumping system, coupled to an autosampler and degasser. Chromatographic separation was performed using a Synergy Hydro-RP column (100 mm × 2 mm, 2.5 μm particle size, Phenomenex). The elution gradient involved a binary solvent system as described previously[29]. Accurate mass data were acquired using a Q Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific), which was equipped with a heated electrospray ionization source operated in negative electrospray mode. Ionization source working parameters were optimized; the heater temperature was set to 300 °C, ion spray voltage was set to 3,500 V. An m/z scan range from 70 to 700 was chosen and the resolution was set at 70,000. The automatic gain control target was set to 1e6 and the maximum injection time was 250 ms. Instrument control and acquisition was achieved using Xcalibur 2.2 software (Thermo Fisher Scientific). All data analysis was conducted using El-MAVEN software[30].

Synthesis of DI-GSH conjugate (N5-(1-((carboxymethyl)amino)-3-((4-methoxy-2-(methoxycarbonyl)-4-oxobutyl)thio)-1-oxopropan-2-yl)glutamine). To a vial charged with dimethyl 2-methylenesuccinate (0.158 g, 1.0 mmol) was added ethanol (1.000 ml) and triethylamine (0.167 ml, 1.200 mmol) and the mixture was cooled in an ice–water bath before the addition of GSH (0.369 g, 1.200 mmol). The resulting suspension was stirred overnight and allowed to slowly warm to room temperature (ice melt) affording a light yellow solution. The mixture was dried under reduced pressure and purified by RP-HPLC as follows: Agilent automated purification system with single quad MS and DAD; Waters AcQuity UPLC I-Class with QDa and UV; XSelect CSH Prep C18 OBD 5 μm 19 × 100 column. Solvents A and B were water with 0.1% formic acid, and acetonitrile, respectively. The method time was 10 min, with a gradient from 10% B to 40% B over 5 min. Samples were loaded at 10% B. The flow rate during the loading was 25 ml min^{-1} and it was raised to 40 ml min^{-1} during separation, affording N^5-(1-((carboxymethyl) amino)-3-((4-methoxy-2-(methoxycarbonyl)-4-oxobutyl)thio)-1-oxopropan-2-yl)glutamine as a white solid (314 mg, 67.4%). ESI m/z (M + H)$^+$ 466.0. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 8.70 (t, J = 6.6 Hz, 1H), 8.32 (d, J = 8.6 Hz, 1H), 8.13 (d, J = 1.2 Hz, 1H), 4.43–4.33 (m, 1H), 3.67 (m, 1H), 3.60 (s, 3H), 3.58 (s, 3H), 3.25 (m, 3H, beneath water peak in DMSO), 3.00–2.70 (m, 3H), 2.70–2.56 (m, 3H), 2.28 (m, 2H), 1.95–1.75 (m, 2H).

Synthesis of Ita-GSH conjugate (2-(((2-(4-amino-4-carboxybutanamido)-3-((carboxymethyl)amino)-3-oxopropyl)thio)methyl)succinic acid). To a vial charged with 2-methylenesuccinic acid (itaconic acid; 0.021 g, 0.163 mmol) was added water (0.651 ml) and GSH (0.05 g, 0.163 mmol). The resulting suspension was heated at 37 °C overnight, affording a pale yellow solution, which was directly purified by RP-HPLC as follows: XSelect Prep C18 5 μm 19 × 100 column. Solvents A and B were water with 0.1% formic acid, and acetonitrile, respectively. The method time was 10 min with a gradient from 5% B to 10% B over 5 min. Samples were loaded at 5% B. The flow rate during the loading and the separation was

$40\,ml\,min^{-1}$. Mass spectral data were acquired from 200–1000 AMU in electrospray positive mode. The product was successfully isolated as 2-(((2-(4-amino-4-carboxybutanamido)-3-((carboxymethyl)amino)-3-oxopropyl)thio)methyl) succinic acid (28.4 mg, 0.065 mmol, 39.9% yield) as a pale white solid. ESI m/z $(M+H)^{+}$ 437.1. ^{1}H NMR (400 MHz, D_2O) δ 4.48 (ddd, $J=8.7, 5.0, 3.3$ Hz, 1H), 3.86 (s, 2H), 3.71 (t, $J=6.3$ Hz, 1H), 2.97 (ddt, $J=12.1, 9.1, 5.6$ Hz, 2H), 2.87–2.52 (m, 5H), 2.43 (td, $J=7.5, 5.4$ Hz, 2H), 2.06 (q, $J=7.3$ Hz, 2H).

Synthesis of $^{13}C_5$-DI ($^{13}C_5$-DI). *E. coli ita23* (provided by the Klamt laboratory)[31] was allowed to grow in ^{12}C-glucose-LB broth (10 ml of $10\,g\,l^{-1}$ bacto-trypsin, $5\,g\,l^{-1}$ yeast extract, $10\,g\,l^{-1}$ NaCl, $0.28\,g\,l^{-1}$ $CaCl_2$, $125\,mg\,l^{-1}$ kanamycin and 0.2% (w/v) of ^{12}C-glucose) overnight at 30 °C and 210 r.p.m., until OD_{420} (the absorbance at 420 nm) reached 2.6. The production of itaconic acid was next initiated by dilution of 100 μl of the above culture into 250 ml of a ^{13}C-glucose minimal media ($5.0\,g\,l^{-1}$ K_2HPO_4, $3.5\,g\,l^{-1}$ KH_2PO_4, $3.5\,g\,l^{-1}$ $(NH_3)NaHPO_4$, $0.25\,g\,l^{-1}$ $MgSO_4$, $11.3\,mg\,l^{-1}$ $CaCl_2$, $1.5\,g\,l^{-1}$ glutamic acid, $0.5\,mg\,l^{-1}$ thiamine, $25\,mg\,l^{-1}$ kanamycin, 1 ml trace element solution and 0.4% ^{13}C-glucose). The trace element solution consists of $1.6\,g\,l^{-1}$ $FeCl_3$, $0.2\,g\,l^{-1}$ $CoCl_2\cdot6H_2O$, $0.1\,g\,l^{-1}$ $CaCl_2$, $0.2\,g\,l^{-1}$ $ZnCl_2\cdot4H_2O$, $0.2\,g\,l^{-1}$ $NaMoO_4$, $0.05\,g\,l^{-1}$ H_3BO_3. The bacteria were allowed to grow for 6 days until the OD_{420} reached 1.9–2.1. The cells were pelleted by centrifugation for 30 min at 14,000g and 4 °C. The supernatant was collected and lyophilized to afford a white powder (3.78 g). The powder was re-dissolved in distilled H_2O (10 ml) and the pH reduced to 2 using concentrated HCl (around 0.5 ml). ^{13}C-labelled itaconate was extracted into ethyl acetate (4×15 ml). The combined organic layers were dried over anhydrous Na_2SO_4 and concentrated in vacuo to afford ^{13}C-labelled itaconate as an off-white solid (283 mg, 1.1 g per litre of culture). ^{1}H NMR (d_6-DMSO, 500 MHz) δ 12.38 (br s, 2H, -COOH), 6.30–5.92 (m, 1H, H-2a), 5.71 (ddd, 1H, $^{1}J_{HC}$: 158.8 Hz, $^{2}J_{HC}$: 12.0 Hz, $^{3}J_{HC}$: 5.9 Hz, H-2b), 3.38–3.03 (m, 2H, H-4); ^{13}C NMR (d_6-DMSO, 125 MHz) δ 172.0 (dt, $^{1}J_{CC}$: 55.7 Hz, $^{2}J_{CC}$: 2.6 Hz, C-1), 167.5 (d, $^{1}J_{CC}$: 68.8 Hz, C-5), 135.3 (tdd, $^{1}J_{CC}$: 69.1 Hz, $^{2}J_{CC}$: 46.7 Hz, $^{2}J_{CC}$: 2.8 Hz, C-3), 127.4 (d, $^{1}J_{CC}$: 70.8 Hz, C-2), 38.1–36.5 (m, C-4) HR-MS $^{13}C_5H_6O_4$ $(M+Na)^{+}$ calcd. 158.0326, found 158.0328.

The reactant ^{13}C-itaconic acid (0.10 g, 9.5 mmol) was dissolved in methanol (1 ml). To this solution was added one drop of concentrated H_2SO_4 and the mixture was refluxed overnight (16 h). The reaction was quenched with saturated $NaHCO_3$ (1 ml), then extracted into dichloromethane (2×2 ml). The combined organic layers were dried over anhydrous Na_2SO_4 to afford a brown liquid. Yield: 97 mg, 80%. ^{1}H NMR ($CDCl_3$, 300 MHz) δ 6.62–6.00 (m, 1H, H-2a), 6.00–5.36 (m, 1H, H-2b), 3.74 (d, 3H, $^{3}J_{HC}$: 3.8 Hz, H-6), 3.67 (d, 2H, $^{3}J_{HC}$: 3.9 Hz, H-7), 3.60–2.99 (m, 2H, H-4); ^{13}C NMR ($CDCl_3$, 75 MHz) δ 172.0–170.6 (m, C-1), 167.5–166.0 (m, C-5), 133.9 (tdd, $^{1}J_{CC}$: 71.9 Hz, $^{2}J_{CC}$ 46.5 Hz, $^{3}J_{CC}$ 3.0 Hz, C-3), 130.0–127.2 (m, C-2), 52.4–52.3 (m, C-6), 52.3–52.2 (m, C-7), 38.7–36.7 (m, C-4) HR-MS $C_2^{13}C_5H_{10}O_4$ $(M+Na)^{+}$ calcd. 168.0639, found 168.0638.

Protein mass spectrometry. BMDMs were plated 2×10^6 cells per well in 6-well plates. Cells were subjected to one of the following: (1) treated with DI (250 μM) for 12 h and then stimulated with LPS (100 ng ml^{-1}) for 1 h, or (2) only stimulated with LPS (100 ng ml^{-1}) for 1 h, or (3) only treated with DI (250 μM) for 12 h, or (4) treated with neither DI nor LPS. Cells were then washed three times with PBS and lysed in 200 μl of urea buffer (8 M urea, 75 mM NaCl, 50 mM Tris pH 8.0, 1 mM EDTA). Lysates were then cleared by centrifugation at 20,000g and protein concentrations were determined by BCA assay (Pierce). 15 μg of total protein per sample were processed further. Disulfide bonds were reduced with 5 mM dithiothreitol and cysteines were subsequently alkylated with 10 mM iodoacetamide. Samples were diluted 1:4 with 50 mM Tris-HCl (pH 8.0) and sequencing grade modified trypsin (Promega) was added in an enzyme-to-substrate ratio of 1:50. After 16 h of digestion, samples were acidified with 1% formic acid (final concentration). Tryptic peptides were desalted on C18 StageTips[32] and evaporated to dryness in a vacuum concentrator. Desalted peptides were labelled with the TMT10plex mass tag labelling reagent according to the manufacturer's instructions (Thermo Scientific) with small modifications. In brief, 0.2 units of TMT10plex reagent was used per 15 μg of sample. Peptides were dissolved in 30 μl of 50 mM HEPES pH 8.5 and the TMT10plex reagent was added in 12.3 μl of MeCN. After 1 h incubation the reaction was stopped with 2.5 μl 5% hydroxylamine for 15 min at 25 °C. Differentially labelled peptides were mixed for each replicate and subsequently desalted on C18 StageTips[32] and evaporated to dryness in a vacuum concentrator.

The peptide mixtures were fractionated by strong cation exchange (SCX) using StageTips as previously described[32] with slight modifications. In brief, one StageTip was prepared per sample by 3 SCX discs (3 M, cat. no. 2251) topped with 2 C18 discs (3 M, cat. no. 2215). The packed StageTips were first washed with 100 μl methanol and then with 100 μl 80% acetonitrile and 0.2% formic acid. Afterwards they were equilibrated by 100 μl 0.2% formic acid and the sample was loaded onto the discs. The sample was transeluted from the C18 discs to the SCX discs by applying 100 μl 80% acetonitrile; 0.2% formic acid, which was followed by three stepwise elutions and collections of the peptide mix from the SCX discs. The first fraction was eluted with 50 μl of 50 mM NH_4AcO, 20% MeCN (pH ~7.2);

the second with 50 μl 50 mM NH_4HCO_3, 20% MeCN (pH ~8.5) and the third with 50 μl 0.1% NH_4OH, 20% MeCN (pH ~9.5). 200 μl of 0.2% acetic acid was added to each of the three fractions and they were subsequently desalted on C18 StageTips as previously described[32] and evaporated to dryness in a vacuum concentrator. Peptides were reconstituted in 10 μl 0.2% formic acid. Both the unfractionated samples and the fractionated, less complex samples were then analysed by LC–MS/MS on a Q Exactive HF as previously described[33,34]. In brief, around 1 μg of total peptides were analysed on an Eksigent nanoLC-415 HPLC system (Sciex) coupled via a 25-cm C18 column (inner diameter of 100 μm, packed in-house with 2.4 μm ReproSil-Pur C18-AQ medium, Dr. Maisch GmbH) to a benchtop Orbitrap Q Exactive HF mass spectrometer (Thermo Fisher Scientific). Peptides were separated at a flow rate of 200 nl min^{-1} with a linear 206-min gradient from 2% to 25% solvent B (100% acetonitrile, 0.1% formic acid), followed by a linear 5-min gradient from 25 to 85% solvent B. Each sample was run for 270 min, including sample loading and column equilibration times. Data were acquired in data-dependent mode using Xcalibur 2.8 software. MS1 spectra were measured with a resolution of 60,000, an automatic gain control target of 3e6 and a mass range from 375 to 2000 m/z. Up to 15 MS2 spectra per duty cycle were triggered at a resolution of 60,000, an automatic gain control target of 2e5, an isolation window of 1.6 m/z and a normalized collision energy of 36.

All raw data were analysed with MaxQuant software[35] version 1.6.0.16 using a UniProt *Mus musculus* database (downloaded on 16 May 2017), and MS/MS searches were performed with the following parameters: The five mass spec runs were grouped together. TMT11plex labelling on the MS2 level, oxidation of methionine, deamidation of asparagine and protein N-terminal acetylation as variable modifications; carbamidomethylation as fixed modification; Trypsin/P as the digestion enzyme; precursor ion mass tolerances of 20 p.p.m. for the first search (used for nonlinear mass re-calibration) and 4.5 p.p.m. for the main search, and a fragment ion mass tolerance of 20 p.p.m. For identification, we applied a maximum false discovery rate of 1% separately on protein and peptide level. We required one or more unique/razor peptides for protein identification and a ratio count for each of the 10 TMT channels. This gave us a total of 4,123 quantified protein groups.

Finally, we normalized the MaxQuant-generated corrected TMT intensities such that at each condition or time point the corrected TMT intensity values added up to exactly 1,000,000, therefore each protein group value can be regarded as a normalized microshare (we did this separately for each TMT channel for all proteins that made our filter cutoff in all the TMT channels. After that we added a pseudocount of 1 to each intensity value in order to account for the noise level and make our fold-change calls more robust for small intensity values. Finally, we $\log_2$-transformed all values and took the average of the $\log_2$ values for all replicates per condition (if replicate samples were present).

Metabolic labelling of nascent protein synthesis with click chemistry. BMDMs were grown in 12-well plates, 10^6 cells per well and treated with DI (250 μM) for 10 h. The cells were then washed three times with methionine-deficient media and incubated for 1 h without methionine in the presence of 200 μM DI. After that, L-azidohomoalanine (Click-iT AHA, C10102, Invitrogen) was added directly to cell media to a final concentration of 50 μM. The cells were then stimulated with LPS for 1 h. In some samples, cells were treated with puromycin (5 μg ml^{-1}) for 2 h before LPS stimulation to block translation. Cells were lysed in 100 μl of lysis buffer (50 mM Tris-HCl, pH 8, 1% SDS) supplemented with protease inhibitor cocktail, phenylmethylsulfonyl fluoride and Na_3VO_4 (Santa Cruz Biotechnology). Lysates were incubated on ice for 30 min, sonicated and cleared by centrifugation at 13,000g for 5 min at 4 °C. Total protein concentration was determined using an RC/DC Protein Assay (Bio-Rad). 30 μg of protein was used for the downstream reaction with 40 nM biotin–alkyne (B10185, Invitrogen) and the reaction was carried out in Click-iT Protein Reaction Buffer Kit (C10276, Invitrogen) according to the manufacturer's protocol. Proteins were separated on 4–20% polyacrylamide gradient gels (Bio-Rad) and biotinylated proteins were detected by western blot with a streptavidin–HRP conjugate (1:1000, #554066, BD Pharmingen). The membrane was stripped using 0.2 M NaOH and reprobed to detect IκBζ (see 'Western blots').

Glutathione measurement. The total GSH concentration in cells was determined by a GSH/GSSH Ratio Detection Assay Kit (Abcam) according to the manufacturer's protocol. In brief, 10^6 BMDMs were lysed in 100 μl of 0.5% NP-40 in PBS, pH 6. Samples were deproteinized using trichloroacetic acid and neutralized by the addition of 1 M $NaHCO_3$ to achieve pH values of 4–6. Collected extracts were diluted with the supplied assay buffer and used directly for GSH measurement.

Cytokine detection. Cytokines in cell supernatants were analysed using DuoSet ELISA kits according to the manufacturer's protocol (R&D Systems). The supernatants from BMDMs were diluted 1:4, those from human blood monocytes were diluted 1:3.

RNA isolation and quantitative real-time PCR. RNA from cultured cells was isolated using a Total RNA I kit (Omega). RNA from the skin of mouse ears was extracted using an RNeasy mini kit (Qiagen) after tissue disruption with sterile zirconium beads on a MagNA Lyser (Roche). Isolated RNA was reverse-transcribed

using AffinityScript Multi-Temp reverse transcriptase (Agilent Technologies) according to the manufacturer's protocol. Reactions were performed in 96-well plates using a SYBR Green PCR Master mix (Thermo Fisher Scientific) using a LightCycler 96 or LightCycler 480 (Roche Diagnostics). All assays were performed at least in duplicate, and reaction mixtures in 20-μl volumes (96-plate) or 10-μl volumes (384-plate) were processed under the following cycling conditions: initial 10 min denaturation at 95 °C, followed by 40 cycles at 95 °C for 10 s, 60 °C for 1 min. Threshold cycle values for each sample were determined by automated threshold analysis. Expression levels of all mRNAs were normalized to reference gene *Actb* in BMDMs or to *Rpl19* in mouse keratinocytes and tissue samples; to *ACTB* in human blood monocytes or *RPLP0* in human keratinocytes. The relative increase in the expression level of a gene was normalized to the level of expression in unstimulated control cells in each experiment. Primer pairs used are listed in Supplementary Table 1. IκBζ-dependent genes selected for analysis in macrophages have been published previously[36].

Lentiviral transduction. Mouse *Nfkbiz* 3′ UTR Lenti-reporter-GFP (#MT-m64048) vector or pLenti-UTR-GFP-Blank vector (#m014) were purchased from Applied Biological Materials. For lentiviral production, a transfection mixture of 1.5 ml of Opti-MEM medium (Invitrogen), 18 μg of psPAX2 (gift from D. Trono, Addgene plasmid #12260), 13 μg of pCMV-V-SVG (gift from B. Weinberg, Addgene plasmid #8454)[37], 20 μg of lentiviral construct, and 105 μl of polyethylenimine (1 mg ml^{-1}, 25 kDa, linear form, Polysciences) was used. The mixture was incubated at room temperature for 20 min and added to HEK-293T cells (ATCC) cultured in DMEM supplemented with 10% FBS, 2 mM L-glutamine and 100 U ml^{-1} penicillin–streptomycin in a 150 cm^2 tissue-culture flask. After 48 h, virus-containing medium was filtered through cellulose acetate filters (45 μm pore-size) and used directly for BV2 transduction in the presence of polybrene (8 μg ml^{-1}, Millipore). The next day, the medium was replaced and two days after this the cells were selected with puromycin at 5 μg ml^{-1}.

Flow cytometry. For measurements of reactive oxygen species, cells were treated with DI (250 μM), loaded with 10 μM CM-H$_2$DCFDA (Invitrogen) at room temperature for 30 min in Hank's balanced salt solution (HBSS). After incubation, cells were rinsed with warm HBSS, collected and directly analysed. The mean fluorescent intensity of living cells is shown. For *Nfkbiz* 3′ UTR reporter analysis, the GFP signal was determined in live cells. For analysis of cell viability, cells were stained with propidium iodide (1 μg ml^{-1}) and the percentage of negative cells was plotted. Cells were acquired on CantoII or LSRII flow cytometers (Becton Dickinson), and data were analysed with FlowJo v.9.5.2 software (Tree Star). For the gating strategy, see Supplementary Fig. 2.

Confocal microscopy. BMDMs were seeded at 8-well multitest microscopy slides (MP Biomedicals). Cells were treated with DI (250 μM, 12 h) and then stimulated with LPS (100 ng ml^{-1}, 30 min). The cells were fixed with 3% paraformaldehyde in PBS for 30 min and permeabilized with 0.1% Triton X-100 in PBS for 15 min. After washing with PBS, samples were blocked with 1% BSA for 15 min and subsequently labelled with p65-specific antibody (1:50, #8242, Cell Signaling), followed by AF568-conjugated anti-rabbit secondary antibody (1:500, #A11011, Thermo Fisher Scientific) in PBS containing 1% BSA. After labelling, the cells were washed with PBS and mounted in 50% (w/v) glycerol in PBS, pH 8.5 containing 4′,6-diamidino-2-phenylindole (DAPI; 1 μg ml^{-1}, Sigma) to label nuclei. Images of random fields of view were acquired using a Leica DMi8 confocal microscope (Leica Microsystems) equipped with a HC PL APO 40×/1.3 NA oil immersion objective and exported with LAS AF Lite software (Leica).

Human blood monocytes isolation and treatment. Blood from healthy donors was acquired from leukoreduction chambers supplied by the Mississippi Valley Regional Blood Center. Peripheral blood mononuclear cells from buffy coats were recovered from the Ficoll interface after a 400*g* centrifugation for 30 min. Monocytes were isolated by adherence on a cell-culture dish for 1 h at 37 °C and 5% CO$_2$ in RPMI containing 1% human serum albumin (Albutein). After extensive washes to remove non-adherent cells, monocytes (≥95% purity) were collected, counted and 5 × 10^5 cells were plated per well in 24-well plates. The cells were treated with DI (125 μM unless stated otherwise) for 12 h and stimulated with LPS (100 ng ml^{-1}) for 1 h or as indicated. In some experiments cells were treated with DI in the presence of EtGSH (1 mM, Santa Cruz Biotechnology).

Primary mouse and human keratinocytes. Primary keratinocytes were isolated from C57BL/6 wild-type newborn mice or human foreskins as previously described[38]. 2 × 10^5 cells in 1 ml of KFSM media (Gibco 10725-018) (Ca^{2+} concentrations of 0.05 mM (mouse) and 0.09 mM (human)) were plated in 12-well tissue culture plates. After 2–3 days of cultivation, mouse or human cells were treated with DI for 12 h and then stimulated with mouse recombinant IL-17A (100 ng ml^{-1}; cat. no. 421-ML, R&D Systems) or human recombinant IL-17A (100 ng ml^{-1}; 7955-IL, R&D Systems) for 4 h or as indicated.

IMQ-induced psoriasis. To induce experimental psoriasis[39], imiquimod (IMQ; imiquimod cream 5%, Perrigo. Co.) was applied daily to mice on both ears (~5 mg per ear) for 7 days. For the DI-treated mice, DI was administered via the intraperitoneal route at 20 mg in 500 μl sterile PBS per mouse one day before IMQ application and daily thereafter for seven days. After seven days, mice were euthanized and ears were used for RNA extraction or histological analysis by performing haematoxylin and eosin staining on 7-μm thick sections after paraffin embedding. The average ear thickness in each sample was quantified from images obtained at identical settings using ImageJ[27,28]. Thickness at five equidistant places in each image was quantified and the mean of these values was used to represent results in each mouse.

SDH activity in mouse heart and liver. DI was administered to mice via the intraperitoneal route at 20 mg in 500 μl sterile PBS per mouse, either once per day for a total length of four days (DI daily) or every 2 h, three times in total (DI overdose). For the daily DI administration, the last injection was 6 h before the tissue collection, and in the DI overdose protocol the last injection was 2 h before collection. Mice were euthanized and heart (~50 mg) and liver tissue (~200 mg) were collected, washed in PBS and mitochondria were isolated with Mitochondria Isolation Kit for Tissue (#89801, Thermo Scientific) according to the manufacturer's protocol. For the analysis of mitochondrial purity, cytoplasmic and mitochondrial fractions were diluted with reducing sample buffer and analysed by western blot for the presence of SDH and GAPDH (see 'Western blots'). SDH activity in isolated mitochondria was analysed using an SDH Activity Colorimetric Kit (#MAK197, Sigma). Isolated mitochondria were directly resuspended in SDH assay buffer. In parallel, protein concentration in each sample was determined using RC/DC Protein Assay (Bio-Rad) and activity was normalized to protein concentration.

Statistical analysis. Unless stated otherwise, standard statistical analyses were performed using MS Excel or GraphPad Prism 7. The type and number of replicates and the statistical tests used are described in the figure legends. Exact *P* values are shown where determined. Individual data points are shown, and the mean ± s.e.m. is reported for analyses where $n > 2$, only the mean is reported where $n \leq 2$.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. The raw and processed RNA-seq data have been deposited in the Gene Expression Omnibus with accession numbers GSE102190 and GSE110749. The original mass spectra may be downloaded from MassIVE (http://massive.ucsd.edu) using the identifier MSV000082101. Source data for the graphical representations found in all figures and Extended Data figures are provided. Source data for western blots (uncropped and unprocessed scans with size marker indications) are presented in Supplementary Fig. 1. All other data that support the findings of this study are available from the corresponding author upon reasonable request.

20. Okada, K. et al. The α-glucosidase inhibitor acarbose prevents obesity and simple steatosis in sequestosome 1/A170/p62 deficient mice. *Hepatol. Res.* **39**, 490–500 (2009).
21. Parada, E. et al. The microglial α7-acetylcholine nicotinic receptor is a key element in promoting neuroprotection by inducing heme oxygenase-1 via nuclear factor erythroid-2-related factor 2. *Antioxid. Redox Signal.* **19**, 1135–1148 (2013).
22. Hartman, M. G. et al. Role for activating transcription factor 3 in stress-induced beta-cell apoptosis. *Mol. Cell. Biol.* **24**, 5721–5732 (2004).
23. Vincent, E. E. et al. Mitochondrial phosphoenolpyruvate carboxykinase regulates metabolic adaptation and enables glucose-independent tumor growth. *Mol. Cell* **60**, 195–207 (2015).
24. Sergushichev, A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. Preprint at https://www.biorxiv.org/content/early/2016/06/20/060012 (2016).
25. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
26. McGuire, V. A. et al. Dimethyl fumarate blocks pro-inflammatory cytokine production via inhibition of TLR induced M1 and K63 ubiquitin chain formation. *Sci. Rep.* **6**, 31159 (2016).
27. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
28. Rueden, C. T. et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18**, 529 (2017).
29. Lu, W. et al. Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone Orbitrap mass spectrometer. *Anal. Chem.* **82**, 3212–3221 (2010).
30. Sahil et al. ElucidataInc/ElMaven: El-MAVEN v0.2.2. (2017).
31. Harder, B.-J., Bettenbrock, K. & Klamt, S. Model-based metabolic engineering enables high yield itaconic acid production by *Escherichia coli. Metab. Eng.* **38**, 29–37 (2016).
32. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
33. Cheng, Z. et al. Pervasive, coordinated protein-level changes driven by transcript isoform switching during meiosis. *Cell* **172**, 910–923.e16 (2018).
34. Keshishian, H. et al. Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol. Cell. Proteomics* **14**, 2375–2393 (2015).

35. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

36. Hildebrand, D. G. et al. IκBζ is a transcriptional key regulator of CCL2/MCP-1. *J. Immunol.* **190**, 4812–4820 (2013).

37. Stewart, S. A. et al. Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493–501 (2003).

38. de Guzman Strong, C. et al. A milieu of regulatory elements in the epidermal differentiation complex syntenic block: implications for atopic dermatitis and psoriasis. *Hum. Mol. Genet.* **19**, 1453–1460 (2010).

39. van der Fits, L. et al. Imiquimod-induced psoriasis-like skin inflammation in mice is mediated via the IL-23/IL-17 axis. *J. Immunol.* **182**, 5836–5845 (2009).
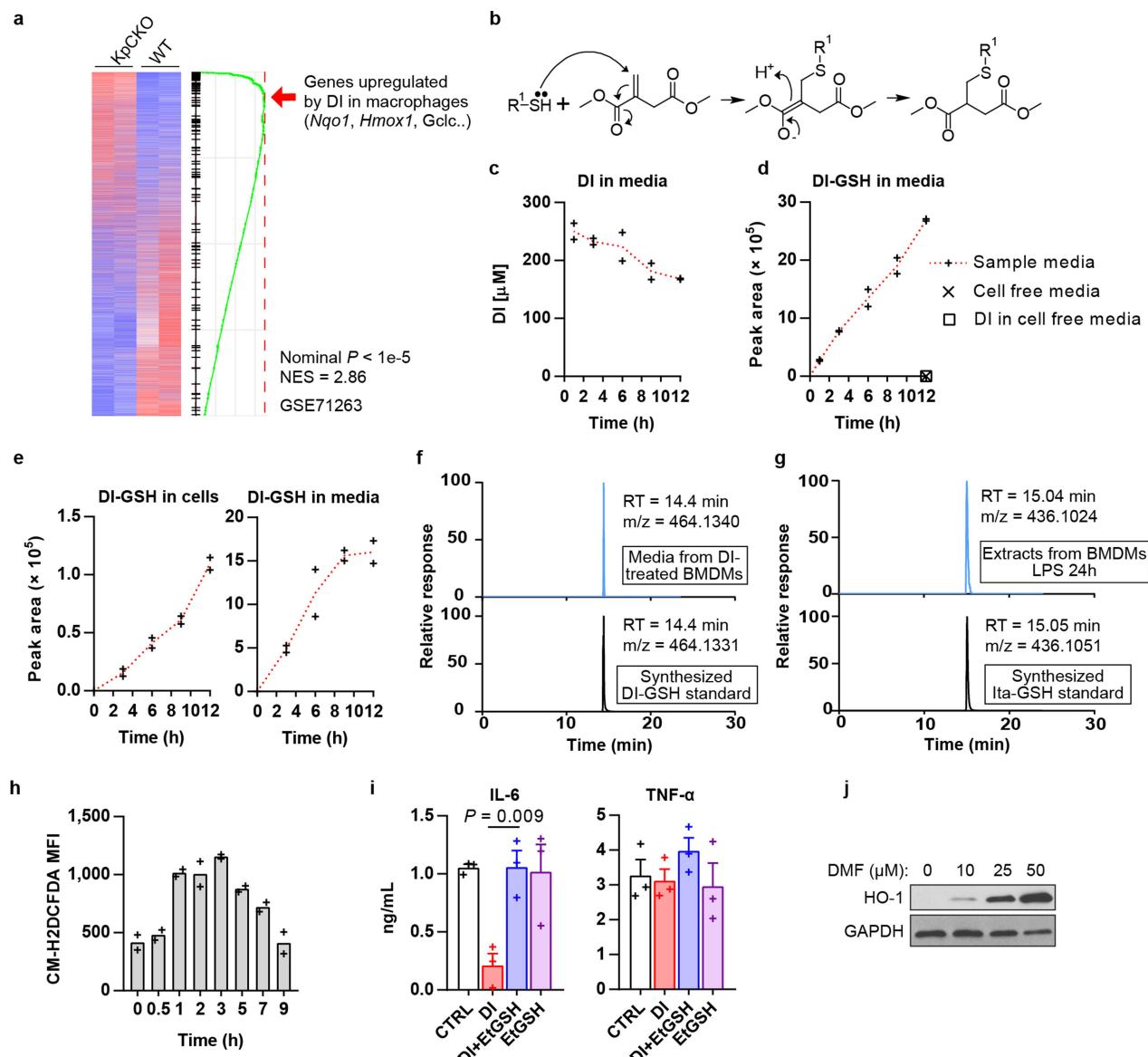
Extended Data Fig. 1 | Detection of DI-GSH and Ita-GSH and electrophilic stress response. a, Transcriptional comparison of KpCKO and wild-type BMDMs and enrichment of the DI gene signature. **b,** The reaction of DI with a thiol group in a Michael reaction. **c,** DI levels in media of BMDMs treated with DI for the indicated time, as determined by GC–MS. Mean of $n = 2$ cultures. **d,** Levels of the DI-GSH conjugate in the media of BMDMs treated with DI for the indicated time, as detected by LC–MS. Mean of $n = 2$ cultures. Data from Fig. 1e are overlaid with data for cell-free media. **e,** Levels of DI-GSH conjugate in BMDMs (left) and in their media (right) after treatment with $^{13}C_5$-labelled DI for the indicated time, as detected by LC–MS. Mean of $n = 2$ cultures. **f, g,** Representative extracted ion chromatograms of DI-GSH detected in the media of BMDMs treated with DI for 6 h compared to the synthesized DI-GSH standard (**f**), and Ita-GSH detected in BMDMs stimulated with LPS for 24 h compared to the synthesized Ita-GSH standard (**g**). $n = 10$ technical replicates. **h,** Detection of reactive oxygen species in BV2 cells treated with DI for the indicated time, as determined by flow cytometry. Mean of $n = 2$ experiments. **i,** Cytokine production in BMDMs treated with DI in the presence of EtGSH and stimulated with LPS for 4 h, mean ± s.e.m., $n = 3$ experiments. **j,** Western blot of HO-1 expression in BMDMs treated with DMF. Representative of three experiments. For gel source data, see Supplementary Fig. 1. Statistical tests used were two-tailed t-tests.
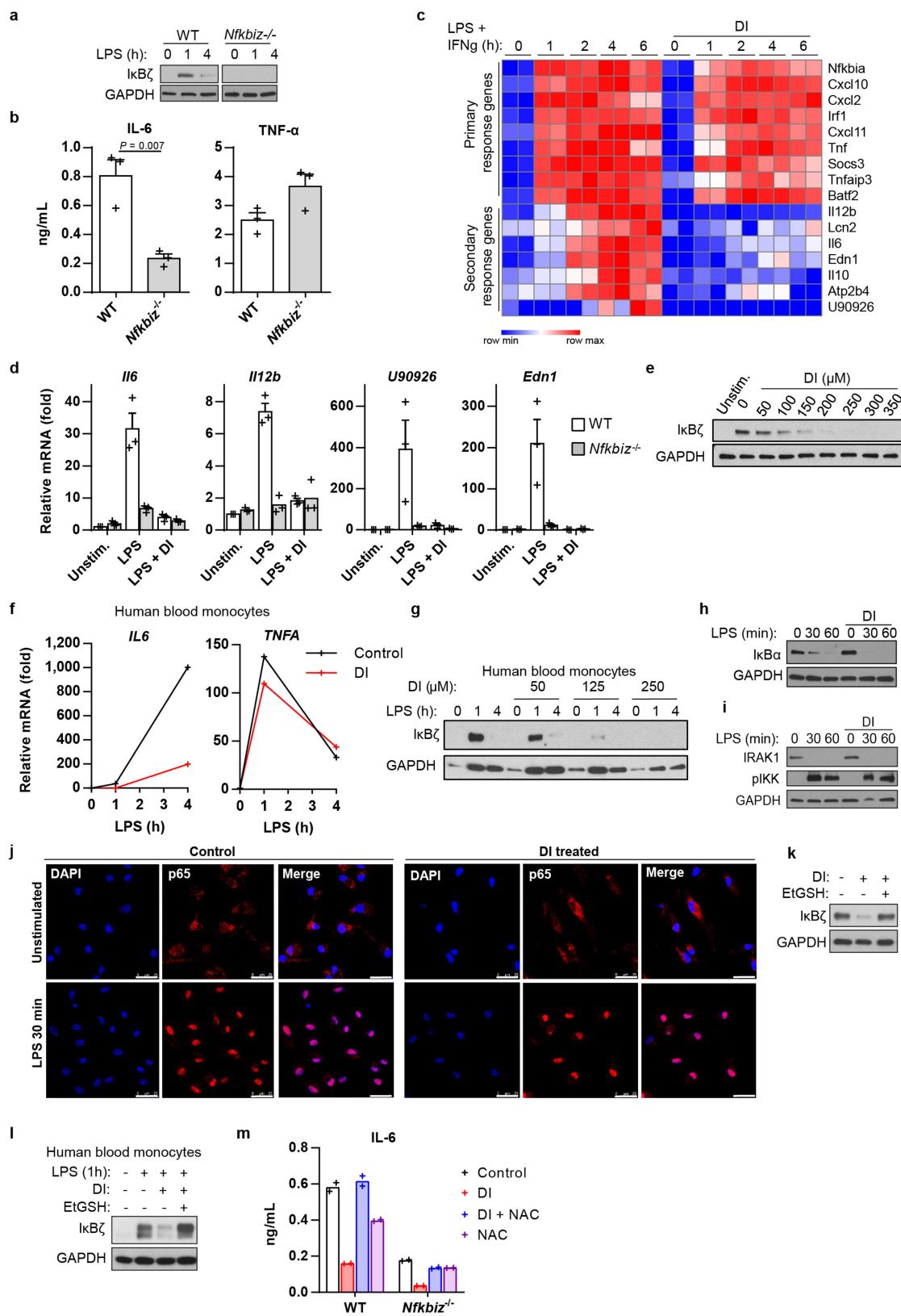
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | DI downregulates secondary transcriptional response to TLR stimulation. a, Western blot of IκBζ expression in wild-type or *Nfkbiz*$^{-/-}$ BMDMs stimulated with LPS. **b**, Cytokine production in wild-type and *Nfkbiz*$^{-/-}$ BMDMs stimulated with LPS for 4 h, mean ± s.e.m., $n = 3$ experiments. **c**, RNA-seq analysis of BMDMs treated with DI and stimulated with LPS and IFNγ. **d**, mRNA expression show the induction of the indicated target genes in wild-type and *Nfkbiz*$^{-/-}$ BMDMs treated with DI and stimulated with LPS for 4 h, mean ± s.e.m., $n = 3$ experiments. **e**, Western blot of IκBζ expression in DI-treated BMDMs stimulated with LPS for 1 h. **f**, mRNA expression in human blood monocytes treated with DI and stimulated with LPS. **g**, Western blot of IκBζ expression in human blood monocytes treated with DI and stimulated with LPS. **h, i**, Western blot of IκBα (**h**) and IRAK1 expression and IKK phosphorylation (**i**) in BMDMs treated with DI and stimulated with LPS. **j**, p65 localization in DI-treated, LPS-stimulated BMDMs. Nuclei are stained with DAPI. Scale bars, 25 μm. Representative of two cultures. **k**, Western blot of IκBζ expression in BMDMs treated with DI in the presence of EtGSH and stimulated with LPS for 1 h. **l**, Western blot of IκBζ expression in human blood monocytes treated with DI in the presence of EtGSH and stimulated with LPS for 1 h. **m**, Cytokine production in wild-type or *Nfkbiz*$^{-/-}$ BMDMs treated with DI in the presence of NAC, stimulated with LPS for 4 h. Mean of $n = 2$ cultures. Representative data from two experiments (**a**), three experiments (**e, h, i, k**), three donors (**f, g**) and two donors (**l**). For gel source data, see Supplementary Fig. 1. Statistical tests used were two-tailed *t*-tests.
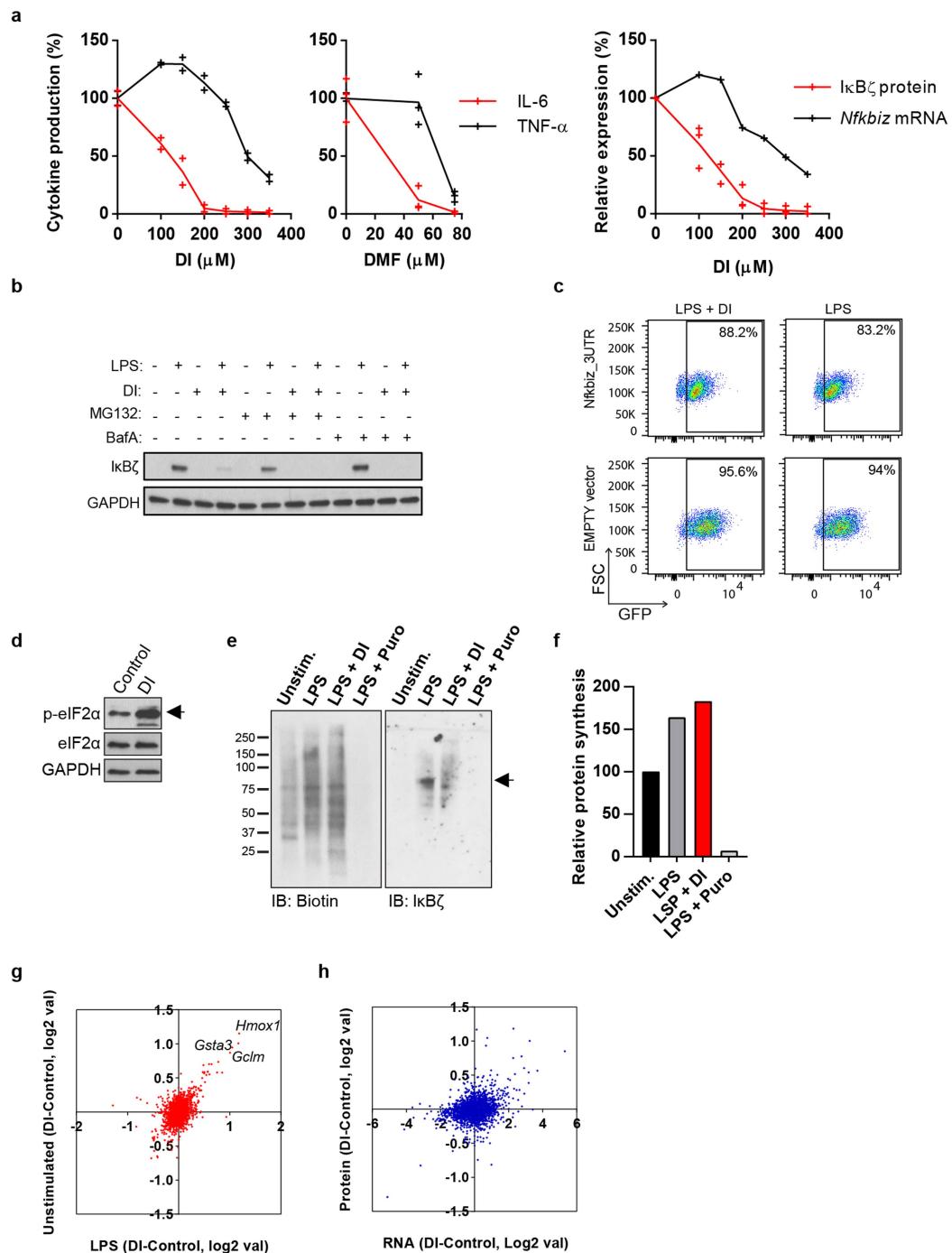
Extended Data Fig. 3 | DI regulates IκBζ at the post-transcriptional level. a, Comparison of the effects of DI on IL-6, TNF and IκBζ on the protein and mRNA levels. Cytokine production is shown in BMDMs treated with DI (left) or DMF (middle) and stimulated with LPS for 4 h (DI), mean of $n = 2$ experiments, or 24 h (DMF), mean ± s.e.m., $n = 3$ experiments. Right, densitometric quantification of IκBζ protein and mRNA expression is shown for BMDMs treated with DI, stimulated with LPS for 1 h. Mean of $n = 3$ experiments, mRNA representative of two experiments. **b**, Western blot of IκBζ expression in BMDMs treated with DI and stimulated with LPS for 1 h. MG132 or bafilomycin A (BafA) were added 30 min before LPS stimulation. **c**, Nfkbiz 3′ UTR reporter expressing GFP in BV2 cells treated with DI (250 μM) for 12 h and stimulated with LPS for 1 h. EMPTY vector expressed GFP only; GFP expression determined by flow cytometry. **d**, Western blot of phosphorylated and total eIF2α in DI-treated BMDMs. **e**, Western blot of nascent protein synthesis detected using biotin–alkyne click chemistry in BMDMs treated with DI and stimulated with LPS for 1 h. The same membrane was reprobed for IκBζ. Representative of two experiments. **f**, Densitometric quantification of the biotin signal in the membrane in **e**. **g**, log fold change of proteomic signal in unstimulated and LPS-stimulated cells. **h**, log fold change of transcript and protein. For **b**–**d**, data is representative of three experiments.
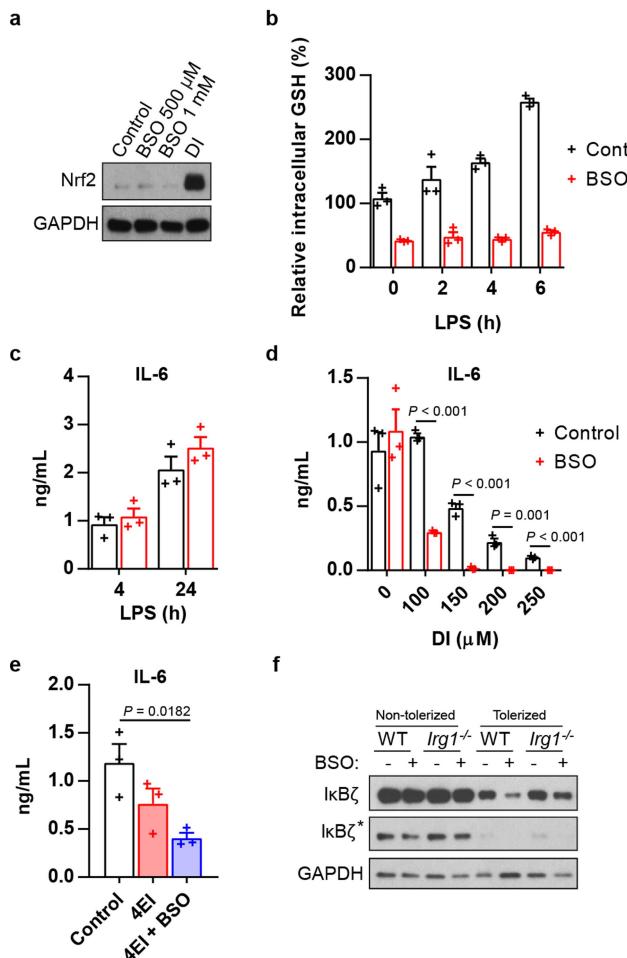
Extended Data Fig. 4 | BSO potentiates the inhibitory effect of DI.
a, Western blot of Nrf2 expression in BMDMs treated with BSO or DI.
b, GSH levels in BMDMs treated with BSO and stimulated with LPS.
Mean ± s.e.m., $n = 3$ cultures. **c**, Cytokine production in BMDMs treated
with BSO and stimulated with LPS. Mean ± s.e.m., $n = 3$ experiments.
d, Cytokine production in BMDMs treated with DI and BSO and
stimulated with LPS for 4 h. Mean ± s.e.m., $n = 3$ experiments. **e**, Cytokine
production in BMDMs treated with 4EI (10 mM) and BSO and stimulated
with LPS for 4 h. Mean ± s.e.m., $n = 3$ experiments. **f**, Western blot of
IκBζ expression in BMDMs tolerized with LPS in the presence of BSO
for 18 h and restimulated for 1 h (see Fig. 2l), asterisk shows the different
exposures. Western blot data are representative of three experiments. For
gel source data, see Supplementary Fig. 1. Statistical tests used were
two-tailed t-tests.
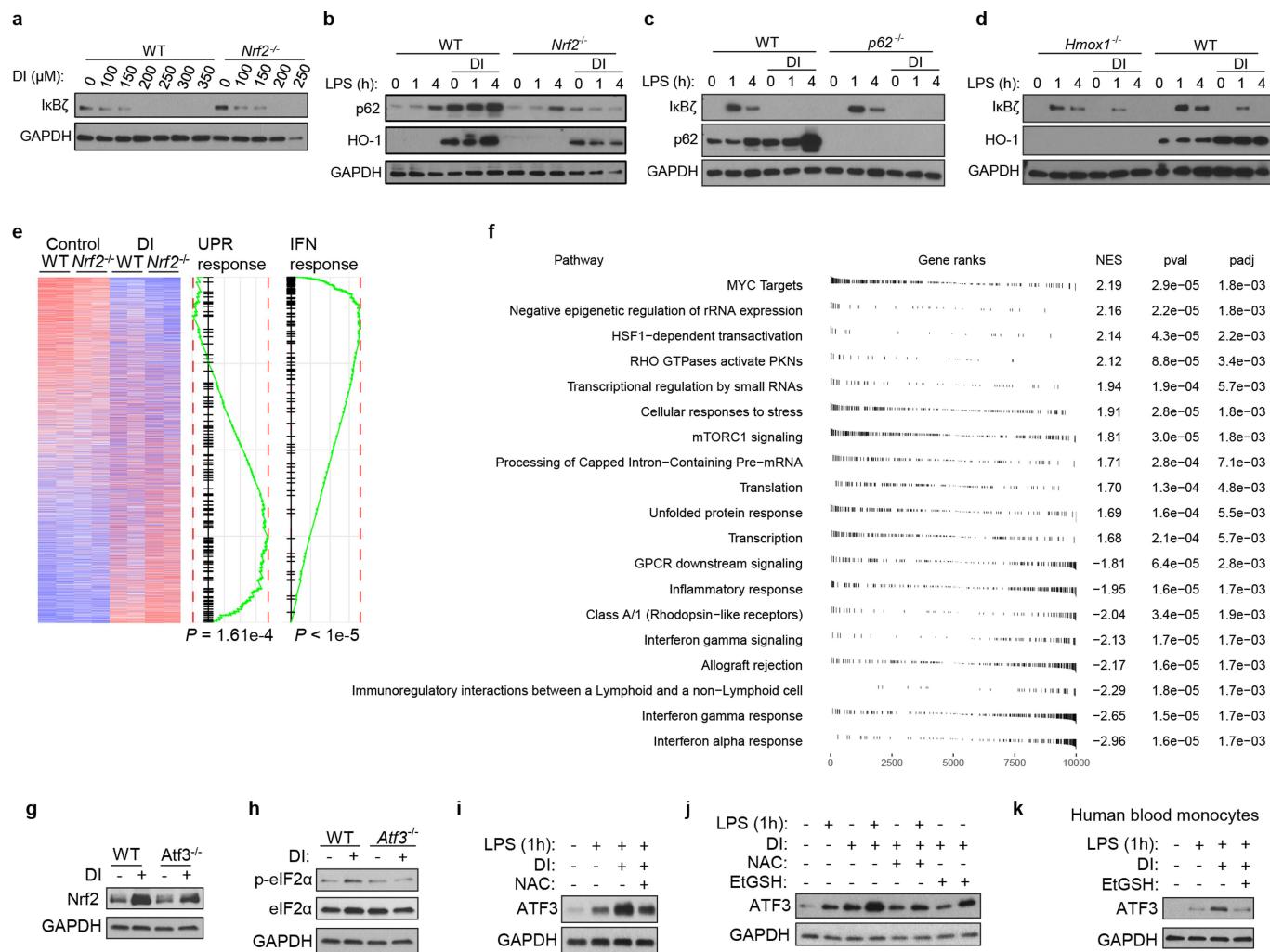
Extended Data Fig. 5 | Nrf2-independent action of DI. a, Western blot of IκBζ expression in wild-type or *Nrf2*[−/−] BMDMs treated with DI and stimulated with LPS for 1 h. **b**, Western blot of p62 and HO-1 in wild-type or *Nrf2*[−/−] BMDMs treated with DI and stimulated with LPS. **c**, Western blot of IκBζ expression in wild-type and p62-deficient BMDMs treated with DI and stimulated with LPS. **d**, Western blot of IκBζ expression in wild-type and *Hmox1*-deficient BMDMs treated with DI and stimulated with LPS. **e**, Transcriptional comparison of *Nrf2*[−/−] and wild-type BMDMs treated with DI and GSEA statistics for unfolded protein response (UPR)

and IFNα pathways. **f**, Pathways regulated by DI in an Nrf2-independent manner. Gene ranks, normalized enrichment score (NES), *P* and adjusted *P* (padj) are shown. **g**, **h**, Western blot of Nrf2 expression (**g**) or phosphorylated and total eIF2α (**h**) in DI-treated wild-type or *Atf3*[−/−] BMDMs. **i–k**, Western blot of ATF3 in BMDMs (**i**, **j**) and human blood monocytes (**k**) treated with DI in combination with NAC or EtGSH and stimulated with LPS. Data are representative of three experiments (**a**, **g**, **i**, **j**), two experiments (**b**, **c**, **h**), one experiment (**d**) and from two donors (**k**). For gel source data, see Supplementary Fig. 1.
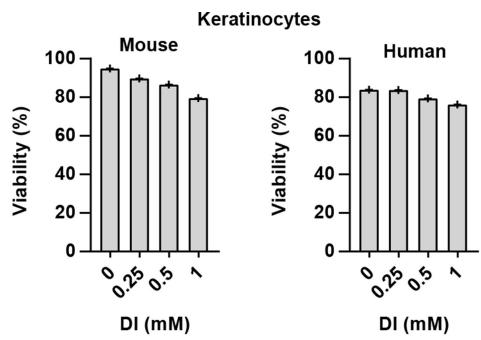
Extended Data Fig. 6 | Viability of keratinocytes after DI treatment.
Mouse and human primary keratinocytes were treated with DI for 12 h and viability was determined by propidium iodide staining and flow cytometry. Percentage of propidium iodide-negative cells is shown. Representative of two mice or donors.

a

DI daily

b

c
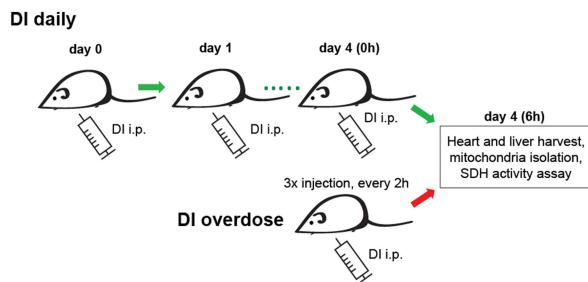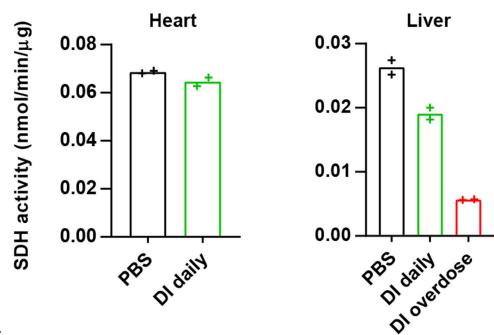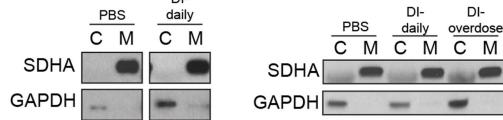
Extended Data Fig. 7 | DI shows a lack of in vivo toxicity. a, Schematic of DI administration for the analysis of succinate dehydrogenase (SDH) activity in the heart and the liver. **b**, SDH activity in the heart and the liver of mice treated as in **a**. Mean of $n = 2$ technical replicates. Representative data from two mice. **c**, Western blot of SDH and GAPDH in mitochondrial and cytoplasmic fractions from the heart and the liver of mice treated as in **a**. Representative of two mice. For gel source data, see Supplementary Fig. 1.

LETTER

https://doi.org/10.1038/s41586-018-0049-7

Genetic identification of leptin neural circuits in energy and glucose homeostases

Jie Xu[1,4], Christopher L. Bartolome[1,2,4], Cho Shing Low[1,3], Xinchi Yi[1], Cheng-Hao Chien[1], Peng Wang[1] & Dong Kong[1,2,3]*

Leptin, a hormone produced in white adipose tissue, acts in the brain to communicate fuel status, suppress appetite following a meal, promote energy expenditure and maintain blood glucose stability[1,2]. Dysregulation of leptin or its receptors (LEPR) results in severe obesity and diabetes[3–5]. Although intensive studies on leptin have transformed obesity and diabetes research[2,6], clinical applications of the molecule are still limited[7], at least in part owing to the complexity and our incomplete understanding of the underlying neural circuits. The hypothalamic neurons that express agouti-related peptide (AGRP) and pro-opiomelanocortin (POMC) have been hypothesized to be the main first-order, leptin-responsive neurons. Selective deletion of LEPR in these neurons with the Cre–loxP system, however, has previously failed to recapitulate, or only marginally recapitulated, the obesity and diabetes that are seen in LEPR-deficient _Lepr^{db/db}_ mice, suggesting that AGRP or POMC neurons are not directly required for the effects of leptin in vivo[8–10]. The primary neural targets of leptin are therefore still unclear. Here we conduct a systematic, unbiased survey of leptin-responsive neurons in streptozotocin-induced diabetic mice and exploit CRISPR–Cas9-mediated genetic ablation of LEPR in vivo. Unexpectedly, we find that AGRP neurons but not POMC neurons are required for the primary action of leptin to regulate both energy balance and glucose homeostasis. Leptin deficiency disinhibits AGRP neurons, and chemogenetic inhibition of these neurons reverses both diabetic hyperphagia and hyperglycaemia. In sharp contrast to previous studies, we show that CRISPR-mediated deletion of LEPR in AGRP neurons causes severe obesity and diabetes, faithfully replicating the phenotype of _Lepr^{db/db}_ mice. We also uncover divergent mechanisms of acute and chronic inhibition of AGRP neurons by leptin (presynaptic potentiation of GABA (γ-aminobutyric acid) neurotransmission and postsynaptic activation of ATP-sensitive potassium channels, respectively). Our findings identify the underlying basis of the neurobiological effects of leptin and associated metabolic disorders.

Leptin deficiency develops secondary to body-weight loss in mice treated with streptozotocin (STZ), a chemical agent that is selectively toxic to pancreatic β-cells and causes loss of insulin (Extended Data Fig. 1a–g). The catabolic consequences and hyperglycaemia are reversed by centrally administered leptin[11,12]. However, as the synthesis and deposition of fat requires insulin, these mice do not gain weight despite the lack of leptin (Extended Data Fig. 1d). We hypothesized that STZ-treated diabetic mice retain leptin sensitivity and can therefore provide an alternative animal model in which to profile leptin-responsive neurons, exempted from the secondary effects of obesity. We systematically analysed two surrogate markers of neuronal activity— FOS, an immediate early gene product, and ribosome protein S6, which becomes phosphorylated (pS6) in excited neurons[13]. Changes in the expression of FOS were largely paralleled by changes in levels of pS6 in the brains of STZ-treated mice; both increased in 53 brain regions and decreased in 10 brain regions (Extended Data Table 1), including in discrete nuclei in the hypothalamus that become enriched in

Lepr-expressing, leptin-responsive neurons (Fig. 1a, b, Extended Data Fig. 1h–j). In addition, similar expression patterns were seen in the brains of non-obese diabetic (NOD) mice (Extended Data Fig. 1m–r), in which diabetes is caused by insulitis[14].

To identify neurons primarily affected by leptin deficiency, we administered leptin in STZ-treated mice. The molecular changes in neuronal activity in STZ-treated mice were broadly corrected 24 h after leptin infusion, whereas leptin treatment for as little as 3 h reversed only the changes in the arcuate nucleus (ARC), and not those in the adjacent lateral hypothalamus (LH) or the dorsomedial hypothalamus (DMH) (Fig. 1a, b; Extended Data Fig. 1i–l), indicating that the ARC contains the majority of neurons that respond directly to leptin stimulation.

We hypothesized that neurons that express FOS and pS6 in the ARC, at least in part, are AGRP neurons, given that STZ induces: 1) the loss of insulin and leptin (Extended Data Fig. 1c, g), both of which inhibit AGRP neurons[15–17]; 2) increased expression of _Agrp_ and _Npy_ mRNA in the mediobasal hypothalamus (Extended Data Fig. 2a); and 3) severe diabetic hyperphagia during both light and dark cycles (Extended Data Fig. 2b–f, Supplementary Video 1), mirroring the voracious feeding in mice with activated AGRP neurons[18,19]. To investigate this, we treated _Npy-hrGFP_ transgenic mice on the basis that they faithfully co-express AGRP and neuropeptide Y (NPY) in the ARC[20]. Indeed, AGRP neurons were remarkably activated, as judged by intensive expression of FOS and pS6, depolarization, and increased firing rates (Fig. 1c–e, Extended Data Fig. 2g–i). To determine the pathological relevance of activated AGRP neurons, we injected an adeno-associated virus vector (AAV) carrying a Cre-dependent hM4Di-mCherry transgene[19] bilaterally into the ARC of _Agrp-IRES-cre_ mice, followed by STZ treatment (Extended Data Fig. 2j). Stimulation of hM4Di with clozapine-_N_-oxide (CNO), thereby inhibiting AGRP neurons, attenuated FOS expression in the ARC (Fig. 1g, h), suppressed diabetic hyperphagia (Fig. 1i, Extended Data Fig. 2k, l) and notably reduced hyperglycaemia (Fig. 1j, k, Extended Data Fig. 2m). CNO injection into STZ-treated, AAV-FLEX-mCherry-transduced mice elicited no obvious changes in these parameters (Extended Data Fig. 2n–q). These results demonstrate that AGRP neurons represent the main ARC neurons that are primarily disinhibited by leptin deficiency, and that their enhanced firing rate contributes essentially to both diabetic hyperphagia and hyperglycaemia.

These findings, however, contradict the prevailing view that AGRP neurons are dispensable for the action of leptin in the brain, as selectively disrupting LEPR in AGRP neurons has previously failed to recapitulate either obesity or diabetes in _Lepr^{db/db}_ mice[9]. We then used CRISPR gene-editing technology[21] to re-examine leptin action and to avoid potential compensatory effects[22,23]. We constructed an AAV carrying a single-guide RNA (sgRNA) targeting the mouse _Lepr_ locus and a Cre-dependent mCherry reporter to indicate virus-transduced neurons (AAV-sgLepr, Fig. 2a, Extended Data Fig. 3a). To examine the efficacy of CRISPR-mediated deletion of LEPR, we mated _Agrp-IRES-cre_ mice with Cre-enabled _Rosa26-LSL-Cas9-GFP_-knockin mice[24] to specifically express Cas9 endonuclease in AGRP neurons, and unilaterally injected AAV-sgLepr into the ARC of

---
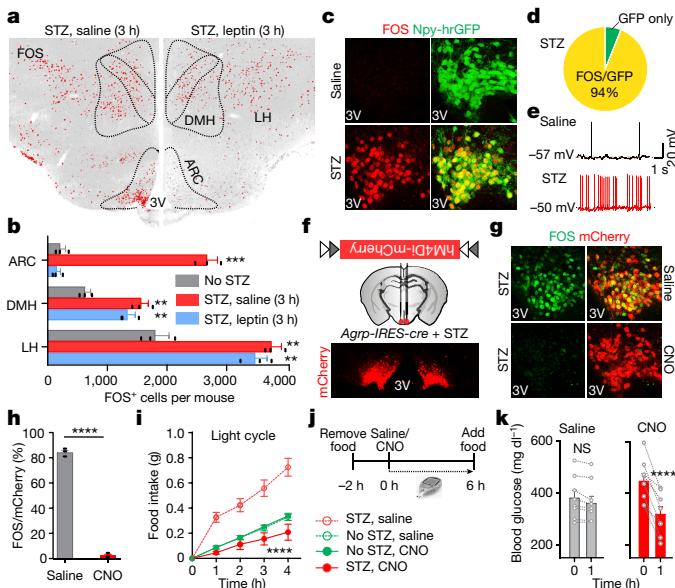
Fig. 1 | Overactivation of AGRP neurons induces diabetic hyperphagia and hyperglycaemia. a, b, FOS immunostaining in the mediobasal hypothalamus of STZ-treated C57BL/6 mice 3 h after the administration of saline or leptin ($n = 3$ mice per group). ARC, arcuate nucleus; DMH, dorsomedial hypothalamus; LH, lateral hypothalamus; 3 V, third ventricle. **c, d**, FOS and hrGFP co-immunostaining in the ARC of *Npy-hrGFP* transgenic mice following saline or STZ treatment ($n = 4$ mice per group). **e**, Representative traces of brain slice whole-cell current-clamp recordings in AGRP neurons of *Npy-hrGFP* mice treated with saline or STZ ($n = 10$ neurons per group from three mice). **f**, AAV pSyn-FLEX-hM4Di-mCherry (top) was injected into the ARC of STZ-treated *Agrp-IRES-cre* mice; the ARC was immunostained for mCherry (bottom). **g, h**, Co-immunostaining of mCherry and FOS in STZ-treated *Agrp-IRES-cre* mice following administration of saline or CNO ($n = 4$ mice per group). **i**, Food intake during light cycle (between 10:00 and 14:00) in ad libitum-fed, virus-transduced *Agrp-IRES-cre* mice following the administration of saline or CNO ($n = 8$ mice per group). **j, k**, Experimental design (**j**) and blood glucose before and 1 h after saline or CNO treatment (**k**) in STZ-treated, virus-transduced *Agrp-IRES-cre* mice ($n = 8$ per group). Data are mean ± s.e.m. and representative of three independent experiments; **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired t-test (**b, h**), paired t-test (**k**) or two-way ANOVA (**i**).

Agrp-IRES-cre::LSL-Cas9-GFP offspring (Extended Data Fig. 3b). We found that reduced expression of *Lepr* mRNA, attenuated leptin-induced phosphorylation of STAT3 (pSTAT3, a marker for LEPR activity) and increased FOS (a marker to indicate disinhibition of neurons) were restricted to the virus-transduced ARC, but not in the contralateral ARC without AAV, nor in the adjacent DMH, which lacked Cre activity (Extended Data Fig. 3c–h), demonstrating effective, neuron-type-specific disruption of LEPR with CRISPR.

To assess the functional relevance of LEPR in AGRP neurons, we injected AAV-sgLepr bilaterally into the ARC of *Agrp-IRES-cre* and *Agrp-IRES-cre::LSL-Cas9-GFP* littermates (Fig. 2b). Extensive co-localization of Cre-enabled mCherry and leptin-induced pSTAT3 was largely diminished in virus-transduced AGRP neurons expressing Cas9 (Fig. 2c, d). Notably, CRISPR-mediated deletion of LEPR in AGRP neurons induced severe hyperleptinaemia, obesity and diabetes, as evidenced by increased fat mass, body weight and daily food intake, reduced energy expenditure and brown adipose tissue activity, elevated serum levels of leptin, insulin and blood glucose, impaired glucose tolerance and extreme insulin resistance in both male and female mice (Fig. 2e–l, Extended Data Fig. 3i–m). These effects are not a result of CRISPR-mediated off-site mutagenesis, as AGRP neuron-specific expression of a CRISPR-insensitive *Lepr* transgene (ci*Lepr*) encoding the long-form LEPR completely prevented such changes in body weight, feeding and blood glucose (Extended Data Fig. 4a–k). As an additional
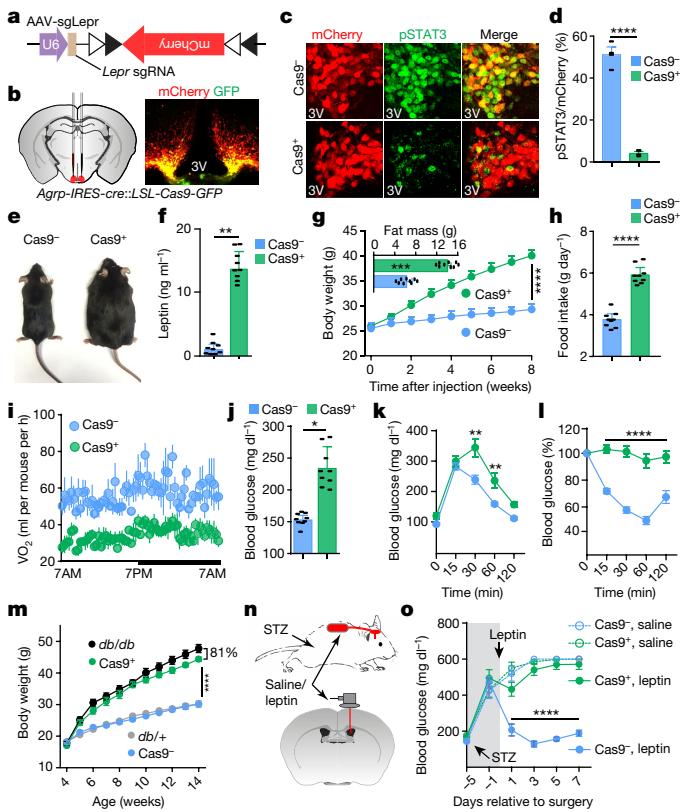
Fig. 2 | CRISPR-mediated deletion of leptin receptors in AGRP neurons results in severe obesity and diabetes. a, b, AAV pU6-sgRNA^{Lepr}::pEF1α-FLEX-mCherry (AAV-sgLepr) (**a**) was injected bilaterally into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice (**b**, left), which was stained for mCherry and GFP (**b**, right). **c, d**, Co-immunostaining of mCherry and leptin-induced pSTAT3 (Tyr705) in ARC of *Agrp-IRES-cre* (Cas9$^-$; **c**, top) and *Agrp-IRES-cre::LSL-Cas9-GFP* mice (Cas9$^+$; **c**, bottom) injected with AAV-sgLepr; quantified in **d** ($n = 4$ mice per group). **e–l**, *Agrp-IRES-cre* (Cas9$^-$) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9$^+$) littermates (**e**), serum leptin levels (**f**), body weight and analysis of fat mass (**g**), daily food intake (**h**), oxygen consumption (**i**), ad libitum-fed blood glucose levels (**j**), glucose-tolerance test (**k**) and insulin-tolerance test (**l**) at eight weeks post-injection with AAV-sgLepr ($n = 9$ mice per group). **m**, Growth curve following AAV-sgLepr injection into the ARC of four-week-old *Agrp-IRES-cre* (Cas9$^-$), *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9$^+$), *Lepr$^{db/+}$* (*db/+*) and *Lepr$^{db/db}$* (*db/db*) mice ($n = 9$ mice per group). **n, o**, Schematic diagram and blood glucose measurement in STZ-treated, virus-transduced *Agrp-IRES-cre* (Cas9$^-$) and *Agrp-IRES-cre::LSL-Cas9-GFP* mice (Cas9$^+$) mice following chronic administration of saline or leptin into the lateral ventricle by implanted osmotic pump ($n = 9$ mice per group). Data are mean ± s.e.m. and representative of three independent experiments; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired t-test (**d, f, g** insert, **h, j**) or two-way ANOVA (**g, k, l, m, o**) with Šidák post hoc test (**k**).

comparison, we injected AAV in parallel into the ARC of *Lepr$^{db/db}$* mice of the same age. Mice with disrupted LEPR in AGRP neurons exhibited around 81% of the weight gain (Fig. 2m), around 85% of the hyperphagia and around 61% of the hyperglycaemia (Extended Data Fig. 3n–p) that is seen in mice that completely lack LEPR (*Lepr$^{db/db}$*), demonstrating that the lack of LEPR in AGRP neurons has a major role in the development of obesity in *Lepr$^{db/db}$* mice. To determine whether LEPR in AGRP neurons is also required for leptin to reverse STZ-induced diabetes, we transduced these neurons with AAV-sgLepr, followed by STZ treatment and implantation of an osmotic pump to chronically infuse leptin centrally (Fig. 2n, Extended Data Fig. 3q). In control diabetic mice, leptin infusion reversed the extreme hyperglycaemia, consistent with previous findings[11]; this effect, however, was abolished in mice expressing Cas9 in AGRP neurons, as was the ability of leptin to restore other catabolic consequences (Fig. 2o, Extended
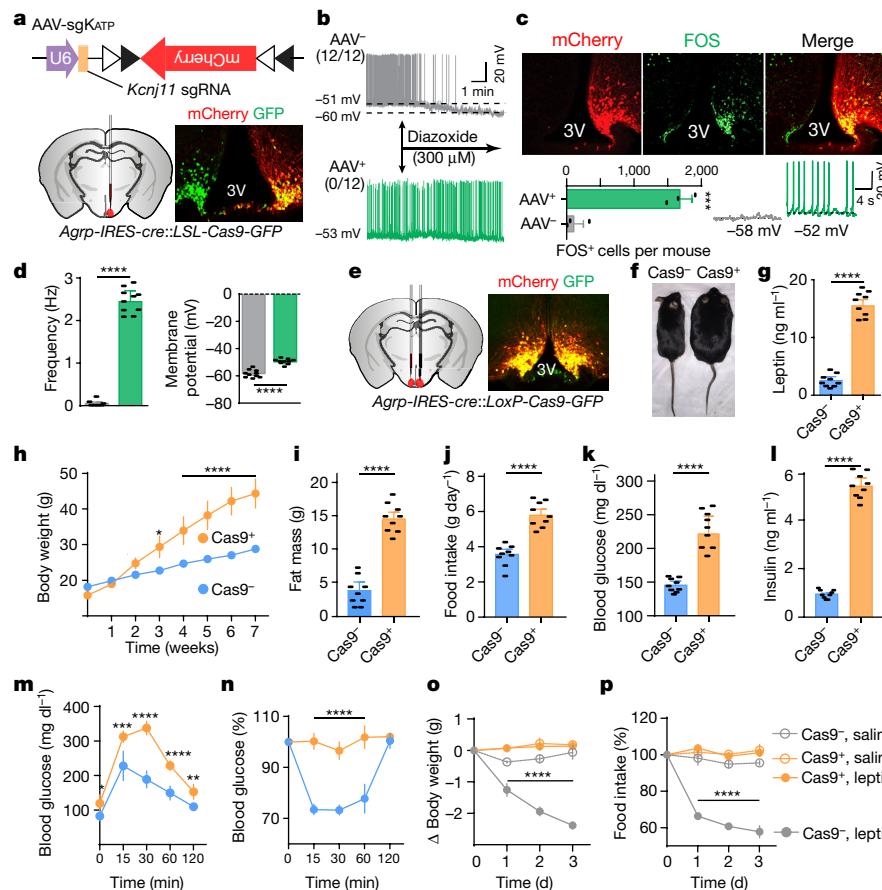
Fig. 3 | K$_{ATP}$ channels in AGRP neurons are required for leptin-mediated regulation of body weight and blood glucose. a, AAV pU6-sgRNA^{Kcnj11}::pEF1α-FLEX-mCherry (AAV-sgK$_{ATP}$) (top) was unilaterally injected into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice, which was co-immunostained for mCherry and GFP. **b**, Representative traces of whole-cell current-clamp recordings in AGRP neurons from non-virus-transduced (grey, AAV$^-$) and AAV-sgK$_{ATP}$-transduced sides (green, AAV$^+$) in 24-h-fasted *Agrp-IRES-cre::LSL-Cas9-GFP* mice. K$_{ATP}$ channel opener diazoxide was applied as indicated ($n = 12$ neurons per group from four mice). **c, d**, mCherry and FOS co-immunostaining and representative traces of brain-slice whole-cell current-clamp recordings (**c**) in AGRP neurons from *Agrp-IRES-cre::LSL-Cas9-GFP* mice that had been injected unilaterally into the ARC with AAV-sgK$_{ATP}$, and quantification of frequency and membrane potential (**d**). $n = 10$ neurons per group from

three mice. **e–n**, AAV-sgK$_{ATP}$ was injected bilaterally into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice, which was co-immunostained for mCherry and GFP (**e**). Representative littermates (**f**), serum leptin levels (**g**), body weight (**h**), fat mass (**i**), daily food intake (**j**), ad libitum-fed state blood glucose levels (**k**), serum insulin levels (**l**), glucose-tolerance test (**m**), and insulin-tolerance test (**n**) at eight weeks post-viral injection. $n = 9$ mice per group. **o, p**, Body weight and daily food intake changes during three-day treatment with leptin versus saline control of *Agrp-IRES-cre* (Cas9$^-$) and *Agrp-IRES-cre::LSL-Cas9-GFP* mice (Cas9$^+$) mice ($n = 9$ mice per group). Data are mean ± s.e.m. and representative of three independent experiments; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired *t*-test (**b, c, g, i–l**) or two-way ANOVA (**o, p**) with Šidák post hoc test (**h, m, n**).

Data Fig. 3r, s). Leptin-induced chronic effects on body weight and food intake were also impaired in animals without STZ treatment (Extended Data Fig. 3t–v). Lastly, we performed CRISPR-mediated deletion of LEPR in ARC POMC neurons and observed no effects either on body weight or on blood glucose, indicating that LEPR in POMC neurons is dispensable (Extended Data Fig. 5a–j). These findings demonstrate that leptin acts primarily on AGRP neurons to maintain energy and glucose homeostases, preventing both obesity and diabetes.

How does leptin inhibit AGRP neurons? Leptin has been shown to hyperpolarize hypothalamic neurons by opening ATP-sensitive potassium channels (K$_{ATP}$) in acute rat brain slices[25]; however, global deletion of the K$_{ATP}$ channel subunits Kir6.2 (also known as KCNJ11) or SUR1 in mice results in only slightly impaired glucose tolerance, without notable changes in body weight[26,27]. To interrogate the functional relevance of K$_{ATP}$ in AGRP neurons, we constructed an AAV carrying an sgRNA targeting the mouse *Kcnj11* locus and a Cre-dependent mCherry transgene (AAV-sgK$_{ATP}$) to achieve CRISPR-mediated deletion of the pore-forming subunit Kir6.2 (Fig. 3a). Following unilateral injection of AAV-sgK$_{ATP}$ into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice, virus-transduced AGRP neurons exhibited increased FOS expression, depolarization,

and firing rates in ad libitum-fed mice, and no responses to a K$_{ATP}$ opener, diazoxide, which induced hyperpolarization and reduced firing of AGRP neurons in fasted animals, compared to the contralateral control neurons (Fig. 3b–d), suggesting functional disruption of K$^+$ efflux. Leptin-induced phosphorylation of STAT3 was not obviously affected (Extended Data Fig. 6a). Following bilateral AAV injection, we observed severe hyperleptinaemia, obesity and diabetes in mice expressing Cas9 in AGRP neurons (Fig. 3e–n), comparable to the phenotypes observed in mice following AGRP-neuron-specific LEPR disruption (Fig. 2e–l). Expression of a CRISPR-insensitive *Kcnj11* transgene (ci*Kcnj11*) in AGRP neurons following CRISPR-mediated deletion of *Kcnj11* prevented the changes in body weight, feeding and blood glucose, excluding possible effects of CRISPR-mediated off-site mutagenesis (Extended Data Fig. 6c–k). To determine whether K$_{ATP}$ channels in AGRP neurons are required for the actions of leptin, we monitored changes in body weight and food intake following three consecutive days of leptin or saline injection (Extended Data Fig. 6b). Whereas leptin markedly reduced body weight and food intake in the control group, it had no effects on mice with disrupted K$_{ATP}$ channels in AGRP neurons (Fig. 3o, p). These results demonstrate that K$_{ATP}$ channels negatively regulate AGRP neurons, and that their absence in

Fig. 4 | Presynaptic GABAergic afferents to AGRP neurons mediate acute leptin inhibition of feeding. a, Representative traces and frequency of sIPSCs recorded with or without leptin incubation in AGRP neurons of 24-h-fasted *Npy-hrGFP* mice. Picrotoxin (PTX) was added to verify GABAergic chloride efflux ($n = 7$ neurons from three mice). **b, c**, AAV pU6-sgRNA$^{Gabrb1/2/3}$::pEF1α-FLEX-mCherry was injected unilaterally into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice, which was stained for mCherry and GFP (**b**); representative traces and quantification of frequency and amplitude of sIPSCs recorded in AGRP neurons from the virus-injected (red) or non-injected sides (green) (**c**; $n = 15$ neurons per group from three mice). **d**, Post-fast refeeding study in virus-transduced *Agrp-IRES-cre* (Cas9$^-$) and *Agrp-IRES-cre::LSL-Cas9-GFP* mice (Cas9$^+$) mice, following administration of saline or leptin ($n = 7$ mice per group). **e**, Representative traces and quantification of electrically evoked IPSCs (eIPSC) paired-pulse ratio in AGRP neurons from 24-h-fasted *Npy-hrGFP* mice ($n = 7$ neurons per group from three mice). **f, g**, Cre-dependent

AAV pSyn-FLEX-hM4Di-mCherry (**f**, top) was injected unilaterally into the ventral DMH (vDMH) of *Vgat-IRES-cre::Npy-hrGFP* mice, which was co-immunostained for mCherry and GFP (**f**, bottom); representative traces and quantification of sIPSCs recorded in AGRP neurons upon CNO incubation (**g**, $n = 5$ neurons). **h–j**, AAV pU6-sgRNA^{Lepr}::pEF1α-FLEX-mCherry and Cre-dependent AAV pMeCP2-FLEX-spCas9 were injected unilaterally into the vDMH of *Vgat-IRES-cre* mice to disrupt presynaptic leptin receptors (**h**), which were co-immunostained for mCherry and leptin-induced pSTAT3 (**i, j**; $n = 4$ mice). **k**, Post-fast refeeding study in mice injected bilaterally in the vDMH of dual-virus-transduced *Vgat-IRES-cre* (Cas9$^-$) and *Vgat-IRES-cre::LSL-Cas9-GFP* (Cas9$^+$), following administration of saline or leptin ($n = 7$ mice per group). Data are mean ± s.e.m. and representative of three independent experiments; **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired *t*-test (**c, d, e, j, k**) or Student's two-tailed, paired *t*-test (**g**).

these neurons notably attenuates leptin activity in maintaining energy balance and glucose homeostasis.

Leptin inhibits hunger-induced appetite and suppresses overeating following an acute fasting period. This effect of leptin, however, was unchanged in mice with disrupted K_{ATP} channels in AGRP neurons (Extended Data Fig. 7a), suggesting an alternative mechanism. Since picrotoxin-sensitive spontaneous inhibitory postsynaptic currents (sIPSCs) in AGRP neurons were markedly reduced following fasting when circulating leptin levels were low, and were restored by further incubation with leptin (Fig. 4a, Extended Data Fig. 7b, c), we considered GABAergic afferent modulation of AgRP neurons as an additional option[28]. Owing to the complex composition of GABA$_A$ receptors, to our knowledge it is not possible yet to achieve genetic ablation of total ionotropic GABAergic neurotransmission in neurons[29] (Extended Data Fig. 7d). To investigate the physiological relevance of GABAergic neurotransmission on AGRP neurons, we constructed an AAV carrying three concatenated sgRNAs targeting the mouse loci (*Gabrb1*, *Gabrb2* and *Gabrb3*) that encode the three GABA$_A$ receptor β-subunits, and a Cre-dependent mCherry transgene (AAV-sgGABA$_A$-R) (Fig. 4b, Extended Data Fig. 7e). Following unilateral injection of the AAV into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice, sIPSCs were eliminated from virus-transduced AGRP neurons (Fig. 4b, c), suggesting

rapid, effective disruption of postsynaptic GABA$_A$ receptors by CRISPR. Bilateral deletion of GABA$_A$ receptors in AGRP neurons induced a transient increase in body weight and a short-term increase in daily food intake, both of which disappeared four weeks after AAV injection, probably owing to a compensatory balance established between excitatory and inhibitory synaptic transmission (Extended Data Fig. 8a–e). During the experimental period, leptin did not exhibit acute suppression of hunger-induced appetite in virus-transduced mice expressing Cas9 (Fig. 4d), but its chronic effects on body weight and food intake remained unaffected (Extended Data Fig. 9a, b). This loss of function was prevented by expression of a CRISPR-insensitive *Gabrb3* transgene (ci*Gabrb3*) in AGRP neurons, which restored GABAergic neurotransmission (Extended Data Fig. 9c–k), demonstrating that GABAergic afferents on AGRP neurons are necessary for the acute inhibition of appetite by leptin.

Finally, we investigated modulation by leptin of GABAergic afferents on AGRP neurons. In fasted mice, but not in ad libitum-fed mice, leptin suppressed the paired-pulse ratio of electrically evoked IPSCs (eIP-SCs), suggesting that there is presynaptic potentiation of GABA release (Fig. 4e, Extended Data Fig. 9l). Since some LEPR-expressing neurons in the ventral DMH (vDMH) are GABAergic[10] and monosynaptically innervate AGRP neurons[30] (Extended Data Fig. 9m, n), we injected

AAV-FLEX-hM4Di-mCherry into the vDMH of *Vgat-IRES-cre::Npy-hrGFP* mice to specifically transduce GABAergic neurons in this nucleus (vGAT[vDMH], Fig. 4f). Inhibition of these neurons upon incubation with CNO eliminated the majority of sIPSCs recorded in AGRP neurons (Fig. 4g), suggesting that vGAT[vDMH] neurons represent the dominant source of GABAergic afferents. We simultaneously transduced vGAT[vDMH] neurons with AAV-sgLepr and a Cre-dependent AAV carrying an *spCas9* transgene and effectively deleted LEPR as revealed by the decrease in leptin-induced pSTAT3 (Fig. 4i, j). Residual LEPR activity presumably comes from non-GABAergic neurons in the DMH[10]. Disruption of LEPR in vGAT[vDMH] neurons prevented leptin-mediated inhibition of fasting-induced overeating (Fig. 4k), replicating the results obtained in mice following postsynaptic GABA$_A$ receptor deletion (Fig. 4d). These results indicate that leptin also engages presynaptic potentiation of GABA release to inhibit AGRP neurons and to suppress hunger-induced appetite.

The present study, summarized in Extended Data Fig. 9o, identifies the fundamental component of the neural circuits governing energy and blood glucose regulation and will facilitate future studies seeking therapeutic interventions for obesity and diabetes. The disparity in phenotypes following CRISPR-mediated gene-editing also highlight the need for a careful re-examination of previous conclusions drawn from conventional genetic ablation studies.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0049-7.

1. Friedman, J. 20 years of leptin: leptin at 20: an overview. *J. Endocrinol.* **223**, T1–T8 (2014).
2. Flak, J. N. & Myers, M. G. Jr. Minireview: CNS mechanisms of leptin action. *Mol. Endocrinol.* **30**, 3–12 (2016).
3. Chen, H. et al. Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in *db/db* mice. *Cell* **84**, 491–495 (1996).
4. Zhang, Y. et al. Positional cloning of the mouse obese gene and its human homologue. *Nature* **372**, 425–432 (1994).
5. Farooqi, I. S. et al. Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. *J. Clin. Invest.* **110**, 1093–1103 (2002).
6. Morton, G. J. & Schwartz, M. W. Leptin and the central nervous system control of glucose metabolism. *Physiol. Rev.* **91**, 389–411 (2011).
7. Kelesidis, T., Kelesidis, I., Chou, S. & Mantzoros, C. S. Narrative review: the role of leptin in human physiology: emerging clinical applications. *Ann. Intern. Med.* **152**, 93–100 (2010).
8. Balthasar, N. et al. Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. *Neuron* **42**, 983–991 (2004).
9. van de Wall, E. et al. Collective and individual functions of leptin receptor modulated neurons controlling metabolism and ingestion. *Endocrinology* **149**, 1773–1785 (2008).
10. Vong, L. et al. Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142–154 (2011).
11. Fujikawa, T., Chuang, J.-C., Sakata, I., Ramadori, G. & Coppari, R. Leptin therapy improves insulin-deficient type 1 diabetes by CNS-dependent mechanisms in mice. *Proc. Natl Acad. Sci. USA* **107**, 17391–17396 (2010).
12. Perry, R. J. et al. Leptin reverses diabetes by suppression of the hypothalamic-pituitary-adrenal axis. *Nat. Med.* **20**, 759–763 (2014).
13. Knight, Z. A. et al. Molecular profiling of activated neurons by phosphorylated ribosome capture. *Cell* **151**, 1126–1137 (2012).
14. Delovitch, T. L. & Singh, B. The nonobese diabetic mouse as a model of autoimmune diabetes: immune dysregulation gets the NOD. *Immunity* **7**, 727–738 (1997).
15. Könner, A. C. et al. Insulin action in AGRP-expressing neurons is required for suppression of hepatic glucose production. *Cell Metab.* **5**, 438–449 (2007).
16. van den Top, M., Lee, K., Whyment, A. D., Blanks, A. M. & Spanswick, D. Orexigen-sensitive NPY/AgRP pacemaker neurons in the hypothalamic arcuate nucleus. *Nat. Neurosci.* **7**, 493–494 (2004).
17. Cowley, M. A. et al. Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* **411**, 480–484 (2001).
18. Aponte, Y., Atasoy, D. & Sternson, S. M. AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat. Neurosci.* **14**, 351–355 (2011).
19. Krashes, M. J. et al. Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J. Clin. Invest.* **121**, 1424–1428 (2011).
20. van den Pol, A. N. et al. Neuromedin B and gastrin-releasing peptide excite arcuate nucleus neuropeptide Y neurons in a novel transgenic mouse expressing strong *Renilla* green fluorescent protein in NPY neurons. *J. Neurosci.* **29**, 4622–4639 (2009).
21. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
22. Luquet, S., Perez, F. A., Hnasko, T. S. & Palmiter, R. D. NPY/AgRP neurons are essential for feeding in adult mice but can be ablated in neonates. *Science* **310**, 683–685 (2005).
23. Wu, Q., Clark, M. S. & Palmiter, R. D. Deciphering a neuronal circuit that mediates appetite. *Nature* **483**, 594–597 (2012).
24. Platt, R. J. et al. CRISPR–Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).
25. Spanswick, D., Smith, M. A., Groppi, V. E., Logan, S. D. & Ashford, M. L. Leptin inhibits hypothalamic neurons by activation of ATP-sensitive potassium channels. *Nature* **390**, 521–525 (1997).
26. Seghers, V., Nakazaki, M., DeMayo, F., Aguilar-Bryan, L. & Bryan, J. Sur1 knockout mice. A model for K$_{ATP}$ channel-independent regulation of insulin secretion. *J. Biol. Chem.* **275**, 9270–9277 (2000).
27. Miki, T. et al. Defective insulin secretion and enhanced insulin action in K$_{ATP}$ channel-deficient mice. *Proc. Natl Acad. Sci. USA* **95**, 10402–10406 (1998).
28. Pinto, S. et al. Rapid rewiring of arcuate nucleus feeding circuits by leptin. *Science* **304**, 110–115 (2004).
29. Rudolph, U. & Möhler, H. Analysis of GABA$_A$ receptor function and dissection of the pharmacology of benzodiazepines and general anesthetics through mouse genetics. *Annu. Rev. Pharmacol. Toxicol.* **44**, 475–498 (2004).
30. Garfield, A. S. et al. Dynamic GABAergic afferent modulation of AgRP neurons. *Nat. Neurosci.* **19**, 1628–1635 (2016).

METHODS

Mice. All animal care and procedures were performed in accordance with national and international guidelines and were approved by the Tufts University/Tufts Medical Center Institutional Animal Care and Use Committee (IACUC), in accordance with NIH guidelines. Mice were group housed (2–5 siblings) at 22–24 °C with a 12-h light–dark cycle, and with ad libitum access to a regular chow diet and water. All diets were provided as pellets. Mice were euthanized by CO_2 narcosis. *Agrp-IRES-cre*[31] (Jax, 012899), *Npy-hrGFP*[20] (Jax, 006417), *Pomc-hrGFP*[32] (Jax, 006421), *Pomc-cre*[8] (Jax, 010714) and *Vgat-IRES-cre*[10] (Jax, 016962) mice were previously generated at the BNORC transgenic core and are available at the Jackson Laboratory. *Rosa26-LSL-Cas9-GFP*[24] (Jax Stock No: 024857) knock-in, *NOD*[33] (Jax Stock No: 001976), *Lepr*[db/db] (Jax Stock No: 000642) and C57BL/6 (Jax Stock No: 000664) mice were obtained from the Jackson Laboratory. Males of mouse lines were used for experiments, and some female mice were used for clinically relevant experiments, particularly with the *Agrp-IRES-cre* and *NOD* mouse lines. Following stereotaxic injection to express AAVs, mice were individually housed with ad libitum access to regular chow diet and water. Littermates of the same sex were randomly assigned to either experimental or control groups.

Virus production. Cre-dependent AAV viral vectors were constructed based on pAAV-pEF1α-FLEX-mCherry-WPRE-pA plasmid[34]. For the AAVs to deliver sgRNAs into the hypothalamus, sgRNAs were designed using online CRISPR tools[35,36] (http://crispr.mit.edu/ and http://chopchop.cbu.uib.no/). The pU6-sgRNA-scaffold cassettes for expressing single or concatenated sgRNAs, including those in AAV vectors AAV-pU6-sgRNA[Lepr]::pEF1α-FLEX-mCherry, AAV-pU6-sgRNA[Kcnj11]::pEF1α-FLEX-mCherry, and AAV-pU6-sgRNA[Gabrb1/2/3]::pEF1α-FLEX-mCherry, were constructed using an in-house designed 'snap-ligation' kit, followed by cloning into the MluI site on pAAV-pEF1α-FLEX-mCherry-WPRE-pA plasmid. For the AAVs to express CRISPR-insensitive (ci) cDNAs in Cre-expressing neurons, including AAV-pEF1α-FLEX-ci*Lepr*, AAV-FLEX-ci*Kcnj11*, and AAV-pEF1α-FLEX-ci*Gabrb3*, cDNAs were generated from the total RNAs extracted from the hypothalamus of C57BL/6 mice by reverse transcription. The DNA fragments containing coding DNA sequences of *Lepr* (encoding the long-form leptin receptor, LEPRB), *Kcnj11* (encoding Kir6.2), and *Gabrb3* (encoding GABA_A receptor β3 subunit) were amplified with the following pairs of PCR primers: *Lepr*-fwd: tttaaaaggatttgcagcgg and *Lepr*-rev: atgacaggctctactggaat; *Kcnj11*-fwd: ggtagacttatcccgccgtg and *Kcnj11*-rev: cctaggccaagccagtgtag; *Gabrb3*-fwd: gaagggatgtggggctttgcg and *Gabrb3*-rev: agcccatcacagagaagcca. Second PCR reactions were performed to introduce silent mutations in sgRNA-binding sites with the following primers: ci*Lepr*-fwd: atgatgtgtcagaaattctatgtggtttttgttacactgggaatttctttatgtgatagctgcacttaacctggcatatccaatc tctccctggaaatttaagttgttttgtggaccccccaataccacagacgatagctttctctcacctgctggagccccc and ci*Lepr*-rev: attacacagttaagtcacac; ci*Kcnj*11-fwd: atgctgtcccgaaagggcattatccctga ggaatatgtgctgacccggctggcaggggaccctgcagagccacgatatagaacgagggaacggagggcccgcttcg tgtccaagaaaggcaac and ci*Kcnj*11-rev: tcaggacaaggaatctggag; ci*Gabrb3*-fwd: atgtggg gctttgcgggaggaaggcttttcggcatcttctcggcccccggtgctggtggcggtggtttgctgcgctcagagcgtaaa cgaccccgggaacatgtcctttgtgaaggagacggtcgacaagctgttgaaaggctacgacataagattacgtccaga ttttggggtcccccagtctgcgtggggatg and ci*Gabrb3*-rev: tcagttaacatagtacagccag. Following Sanger DNA sequencing to verify the mutations, the obtained DNA fragments were ligated into the AscI and NheI restriction enzyme sites of pAAV-pEF1α-FLEX-mCherry-WPRE-pA plasmid. The listed AAV vectors were packaged at the Boston Children's Hospital Viral Core. For CRISPR-insensitive GABA_A receptor subunit re-expression experiments following CRISPR-mediated deletion in AGRP neurons, cDNA encoding GABA_A receptor β3 subunit was selected since this subunit is abundantly expressed in AGRP neurons according to a recent gene-profiling study[37]. The construction and generation of AAV-FLEX-spCas9 virus will be reported elsewhere (C.-H. Chieng, unpublished data). AAV-pSyn-FLEX-hM4Di-mCherry[19] virus was generated from University of North Carolina Vector Core. AAV-pEF1α-FLEX-Synaptophysin-mCherry[38] virus was obtained from MIT Viral Gene Transfer Core. In addition, the following were used (AAV coat serotypes, titre viral molecules/ml): pU6-sgRNA[Lepr]::pEF1α-FLEX-mCherry (2/8, 1.9×10^{14}), pU6-sgRNA[Kcnj11]::pEF1α-FLEX-mCherry (2/8, 7.2×10^{14}), pU6-sgRNA[Gabrb1/2/3]::pEF1α-FLEX-mCherry (2/8, 1.7×10^{15}), AAV-FLEX-ci*Lepr* (DJ, 8.5×10^{13}), AAV-FLEX-ci*Kcnj11* (DJ, 2.2×10^{14}), AAV-FLEX-ci*Gabrb3*-GFP (DJ, 1.9×10^{14}), AAV-FLEX-SpCas9 (DJ, 1.8×10^{14}) and AAV-hSyn-FLEX-hM4Di-mCherry (2/8, 6.7×10^{12}). Viral aliquots were stored at −80 °C before stereotaxic injection.

Stereotaxic surgery. Stereotaxic surgery to deliver AAV into the hypothalamus of mice was performed as previously described[39]. In brief, 4–8-week-old mice were anaesthetized with ketamine (75 mg/kg) and xylazine (5 mg/kg) diluted in saline (0.9% NaCl in water) and fixed on a stereotaxic apparatus (KOPF model 922) with ear-bars. After exposing the skull via a small incision, a small hole was drilled for injection based on coordinates to bregma. A pulled-glass pipette with 20–40-μm tip diameter was inserted into the brain and AAV viruses (50–150 nl per injection

site) were injected by an in-house-built air-puff system. A micromanipulator (Grass Technologies, Model S48 Stimulator) was used to control injection speed at 25 nl/min and the pipette was left in position for another 5 min to allow enough absorption and spreading of AAVs before being withdrawn. For postoperative care, mice were injected intraperitoneally with meloxicam (0.5 mg/kg) for two continuous days. All stereotaxic injection sites were verified under electrophysiological microscopy (for electrophysiology-related studies) or by immunohistochemistry (for anatomy and in vivo studies). All 'missed' or 'partial-hit' animals were excluded from data analyses. Animals were allowed to recover from surgery for 1 week and their body weight and health conditions were closely monitored during recovery. Coordinates and injection volume used in the studies are: the ARC (anterior-posterior (AP): −1.40 mm, dorsal-ventral (DV): −5.80 mm, left-right (LR): ± 0.30 mm, 150 nl/side) and the vDMH (AP: −1.80 mm, DV: −5.30 mm, LR: ± 0.30 mm, 50 nl/side).

Brain slice preparation and electrophysiology. Mice less than eight weeks of age were anaesthetized by inhalation of isoflurane. Coronal sections (300-μm-thick) were cut with a Leica VT1000S vibratome and then incubated in carbogen-saturated (95% O_2/5% CO_2) ACSF (in mM: 125 NaCl, 2.5 KCl, 1 $MgCl_2$, 2 $CaCl_2$, 1.25 NaH_2PO_4, 25 $NaHCO_3$, 10 glucose) at 34 °C for 30–45 min before recording. All recordings were obtained within 4 h of slicing at room temperature. Whole-cell recordings were obtained from arcuate AGRP neurons visualized under infrared differential interference contrast (IR-DIC) using patch pipettes with pipette resistance of 2.5–4.5 MΩ. To identify infected AGRP neurons, mCherry or GFP fluorescence or both were detected using epifluorescence illumination. For sIPSC and sEPSC recordings, the internal solution contained (in mM) 135 $CsMeSO_3$, 10 HEPES, 1 EGTA, 3.3 QX-314 (Cl⁻ salt), 4 Mg-ATP, 0.3 Na-GTP, 8 Na_2-phosphocreatine (pH 7.3 adjusted with CsOH; $295 \, mOsm \, kg^{-1}$). For current-clamp recordings, the internal solution consisted of (in mM) 135 $KMeSO_3$, 3 KCl, 10 HEPES, 1 EGTA, 0.1 $CaCl_2$, 4 Mg-ATP, 0.3 Na-GTP, 8 Na_2-phosphocreatine (pH 7.3 adjusted with KOH; $295 \, mOsm \, kg^{-1}$). For K_{ATP} knockout identification, mice were fasted for 24 h to activate AGRP neuronal activities before brain slice preparation. Diazoxide (Sigma D9035) was added to the bath solution (300 μM). For leptin-induced sIPSC recording, mice were fasted for 24 h and 100 nM mouse leptin was added into bath solution. Mouse leptin was purchased from A. Parlow (NHHP, NIDDK) as 1-mg powder aliquots. A stock solution of 78 μM was made by dissolving 1 mg leptin with 500 μl 15 mM HCl and 300 μl 7.5 mM NaOH. When bath incubation for brain slices was used, a 1:780 dilution was prepared to achieve the final concentration of 100 nM. Picrotoxin (100 μM) (PTX, TOCRIS 1128) was added to block ionotropic GABA_A receptors. For inhibitory hM4Di-related assays, clozapine-*N*-oxide (CNO, 10 μM, NIH Drug Supply Program) was added to the bath solution. To assess IPSCs paired-pulse ratio, pairs of electrical microstimulation pulses were delivered at 50-ms inter-pulse intervals and selected low threshold to ensure recovery of the presynaptic terminals (30 s interval). Recordings were made using an Axoclamp 700B amplifier (Axon Instruments) at room temperature. Data were filtered at 3 kHz and sampled at 10 kHz. Series resistance, measured with a 5-mV hyperpolarizing pulse in voltage clamp, was on average under 20 MΩ and less than 25 MΩ, uncompensated. All voltage-clamp recordings were made from cells held at −60 mV. For current-clamp recordings, membrane potentials were corrected for a ~8 mV liquid junction potential.

Immunohistochemistry. Immunohistochemistry was performed as previously described[39]. In brief, mice were transcardially perfused with 10% formalin and the brains were post-fixed for 1–2 days. Brains were sectioned coronally at 40 μm using a Leica microtome (Leica SM2010R). Brain sections were washed in PBS with 0.25% Triton X-100 (PBT, pH 7.4) and incubated in 3% normal donkey serum (Jackson ImmunoResearch Laboratories, 017-000-121) in PBT-azide for 2 h. Slides were then incubated overnight at room temperature in a primary antiserum. After washing in PBS, sections were incubated in fluorescein-conjugated donkey IgG. Primary antibodies used in the current study and their dilutions are: rabbit anti-DsRed (Clontech, 632496; 1:2000), chicken anti-mCherry (EnCor Biotechnology, CPCA-mCherry; 1:2000), chicken anti-GFP (Aves Labs, GFP-1010; 1:2000), rabbit anti-hrGFP (Agilent Technologies, 240142; 1:1000), goat anti-FOS (Santa Cruz Biotechnology, sc-52-g; 1:150), rabbit anti-pSTAT3 (Cell Signaling Technology, 9145 S; 1:1000), rabbit anti-pS6 (Ser235,236) (Cell Signaling Technology, 4858; 1:1000). Secondary antibodies include Alexa Fluor 594–donkey anti-rabbit IgG (Invitrogen, A-21207; 1:200), Alexa Fluor 594–donkey anti-chicken IgG (Jackson ImmunoResearch, 703-585-155; 1:200), Alexa Fluor 488–donkey anti-chicken IgG (Jackson ImmunoResearch, 703-545-155; 1:200), Alexa Fluor 488–donkey anti-rabbit IgG (Invitrogen, A-21206; 1:200), Alexa Fluor 488–donkey anti-goat IgG (Invitrogen, A-11055; 1:200), Alexa Fluor 594–donkey anti-goat IgG (Invitrogen, A-11058; 1:200). For FOS and pS6 staining following leptin treatment, leptin (5 mg/kg) was intraperitoneally injected in STZ-treated mice, and mice were perfused 3 h later. For pSTAT3 staining, mice were fasted overnight for 24 h, followed by 5 mg/kg intraperitoneal leptin injection[40]. Forty-five minutes later, mice were perfused and brains were dissected out. Following PBS washes

$(3 \times 10\,\text{min})$, brain slices were pre-treated with 1% NaOH for 15 min as previously described. Brain sections were mounted onto SuperFrost (Fisher Scientific 22-034-980) slides and then visualized with a Leica TCS SPE Confocal Microscope (Leica Microsystems) using $10\times$, $20\times$ oil immersion, or $63\times$ oil immersion, or Olympus VS120 Virtual Slide Microscope (Olympus) for whole-brain scanning. Images were imported into ImageJ or Fiji (NIH) software for further analysis.

RNA in situ hybridization. RNA in situ hybridization (ISH) was performed using RNAscope 2.5 HD Assay- Brown kit (acdbio, 322300). In brief, mice were transcardially perfused with 4% paraformaldehyde (PFA) and the brains were post-fixed for 1–2 days. Brain samples were sectioned coronally at $15\,\mu\text{m}$ using a Thermo Scientific cryostat microtome (Thermo Scientific, HM 525) and mounted on SuperFrost Plus slides (Fisher Scientific, 12-550-15). Slides were washed for 5 min with PBS and treated with hydrogen peroxide for 10 min. After washing in distilled water, slides were submerged into boiling $1\times$ retrieval solution for 5 min and washed with distilled water. A hydrophobic barrier was then created and the probe against *Lepr* mRNA (Advanced Cell Diagnostics, 402731, Mm-Lepr) was added onto slides. After 2 h incubation at 40 °C, Hybridize-Amp probes #1–6 were applied on slides sequentially to amplify signals. Finally, DAB was added onto slides to visualize signals and a VS120 Virtual Slide Microscope (Olympus) was used for whole-brain scanning.

Food intake, body weight and body composition analysis. Food intake, body weight and body composition studies were performed as previously described[39]. For daily food intake assays, food pellets were weighed at 10:00 each day for four continuous days and an average of three-day food intake was calculated. Light-cycle food intake was measured between 07:00 and 19:00 and dark-cycle food intake was measured between 19:00 and 07:00. For virus-transduced knockout experiments, animals were singly housed for one week after surgery and maintained for two weeks to allow sufficient expression of AAV-expressed transgene. Body weight and food intake were measured weekly, including the week before surgery. For inhibitory hM4Di-related studies, mice were singly housed and food pellets were weighed before and after intraperitoneal injection of CNO (0.3 mg/kg) during light cycle (10:00 to 14:00) or dark cycle (20:00 to 00:00). For studies in NOD mice, body weight and food intake were measured from three months of age, before the majority of diabetes symptoms develop[33]. For *Lepr*$^{db/db}$ mice, food intake and body weight were tracked from four weeks of age. Body fat mass was measured using echoMRI analysis, and oxygen consumption was measured using metabolic chambers of a Comprehensive Laboratory Animal Monitoring System (CLAMS, Columbus Instruments) from the Adipose Tissue Biology and Nutrient Metabolism Core at the Boston Nutrition and Obesity Research Center (BNORC). Mice were acclimatized in the chambers for 48 h before data collection. Mice with 'missed' injections or incomplete 'hits' were excluded from analysis after post hoc examination of mCherry or GFP expression. In this way, all measurements were randomized and experimenters were blind to treatments.

Blood glucose, glucose and insulin tolerance test. Blood glucose concentrations were measured using a blood glucose meter and glucose test strips (OneTouch Ultra). For virus-induced knockout experiments, blood glucose was measured weekly following surgery. For inhibitory DREADDs (designer receptors exclusively activated by designed drugs) studies, food pellets were removed 2 h before CNO injection. Blood glucose was measured at 0, 1, 2, 4 and 6 h without food following CNO treatment. For glucose tolerance tests (GTT), mice were fasted overnight. After the fasting period, mice received an intraperitoneal injection of 20% glucose (1 g/kg) and blood glucose levels were measured at 0, 15, 30, 60 and 120 min time points post-glucose injection. For insulin tolerance tests (ITT), mice were fasted for 4 h and blood glucose levels were measured at 0, 15, 30, 60, and 120 min time points following. intraperitoneal injection of insulin (0.75 U/kg).

STZ treatment. STZ (Sigma-Aldrich Cat#S0130) was injected intraperitoneally in C57BL/6 mice at varying doses from 75 to 150 mg/kg. STZ injection of 125 mg/kg was selected for subsequent experiments to induce gradual onset of hyperglycaemia and characteristic diabetic symptoms. Health conditions were closely monitored twice a day for two days following STZ injection. Three days after injection, blood glucose was measured to assess the extent of STZ-induced diabetes. Other metabolic studies included monitoring of body weight, food intake, micturition analysis, serum insulin and serum leptin levels.

Leptin effects on body weight and food intake. The effects of leptin on body weight and food intake were assessed as described[40]. In brief, mice (6–7 weeks old) were individually housed and acclimated to handling for one week. These mice were then injected intraperitoneally with saline (0.9% NaCl) or leptin (2 mg/kg per injection, at 07:00 and 19:00) every 12 h for 3 days. Body weight and food intake were measured at baseline and on each day on which saline injections or leptin injections were administered. Body weight and food intake were averaged during the three days before injections to obtain baseline values used for calculating per cent changes. To assess leptin's ability to acutely suppress hunger-induced appetite, mice were singly housed and fasted for 24 h. On the second day, a high dose of leptin (5 mg/kg) was injected intraperitoneally in mice at 07:00 and pre-weighed food was placed back in the cage and monitored for the following 24 h.

Osmotic pump. Intracerebroventricular (i.c.v.) cannula minipump implantation was performed as previously described[11]. In brief, a cannula was implanted into the cerebral lateral ventricle (AP: −0.50 mm, ML: ± 1.3 mm, DV: −2.3 mm), and a mini-osmotic pump (model 1007D, Alzet) was implanted subcutaneously via a catheter connected to the cannula for i.c.v. infusion. The mini-osmotic pump was filled with either leptin (454 ng/μl) or sterile saline solution. Food intake, body weight and glucose were measured on days 1, 3, 5 and 7 post-surgery.

Serum leptin and insulin measurement. Tail vein blood was collected for ELISA assays as previously described[11]. In brief, about 20–50 μl tail vein blood was collected from mice that were food restricted for 2 h before blood collection. Blood was collected with heparinized capillary tubes (Fisher Scientific, 22-260-950) and then centrifuged at 3,000g for 20 min to collect serum and stored at −80 °C. For insulin levels, Ultra Sensitive Mouse Insulin ELISA Kit (Crystal Chem, 90080) was used. For leptin levels, Mouse Leptin ELISA Kit (Crystal Chem, 90030) was used.

Reverse transcription and quantitative PCR assay. Brains were dissected and preserved in RNAlater solution (Thermo Fisher Scientific, AM7020). Total RNA was extracted using TRIzol Reagent (Thermo Fisher Scientific, 15596-018) and reverse transcribed using AffinityScript QPCR cDNA Synthesis Kit (Agilent Technologies #600559). TaqMan probes (all from Applied Biosystems, catalogue number in parentheses) used for quantitative PCR (qPCR) include those for *Agrp* (Mm00475829_g1), *Npy* (Mm00445771_m1), and *Pomc* (Mm01323842_m1), *Lepr* (Mm00440181_m1). Expression of the genes was normalized to that of 18 S ribosomal RNA (Applied Biosystems, 4319413E). qPCR was performed on a StepOnePlus Real-Time PCR System (Applied Biosystems).

Behavioural assays and video analysis. Mice were placed in a $25 \times 40 \times 20$-cm arena to assess feeding duration and 1-h food intake during the light phase (10:00−11:00). Food and nest zones were designated at opposite ends of the arena. A food bowl containing standard chow diet was placed. The spatial locations were tracked and analysed using EthoVision XT 10 software (Noldus) and CCD cameras (Motic).

Thermal imaging of brown adipose tissue. Thermal imaging was performed using a calibrated thermal imaging camera attachment (FLIR One). Mice were anaesthetized and the fur on their back covering the interscapular brown adipose tissue (iBAT) area was shaved. After a one-week recovery and acclimation period, freely moving mice were thermally imaged and their subcutaneous iBAT temperatures were calculated accordingly. Thermal signals from the flank area were used as control.

Cell culture and indel analysis. Neuro-2a (N2a) cells (ATCC, CCL131) were cultured in Dulbecco's modified Eagle's medium (Sigma-Aldrich, D8437) supplemented with 10% fetal bovine serum (Thermo Fisher, 16000044), 500 μg/ml penicillin–streptomycin–glutamine (Thermo Fisher, 10378016) at 37 °C with 5% CO_2. Cells were routinely tested for mycoplasma using PCR detection kit (ATCC, 30-1012 K). Indel analysis was performed with the GeneArt Genomic Cleavage Detection Kit (Thermo Fisher, A24372). In brief, the U6-sgLepr cassette was cloned into the pCas9-GFP backbone (Addgene, 44719) following MluI (New England Biolabs, R3198S) digestion. Cells were transfected with Lipofectamine 3000 (Thermo Fisher, L3000008) and GFP fluorescence was used to assess the transfection efficiency after 48 h. Genomic DNA was extracted and PCR amplification was performed using the following primers: *Lepr* on-target primer forward, cttctctggaaggtagacgctc; reverse, gaccttgctcattcccaaag; *Gpr108* off-target forward, tgagagtcagccggtggata; reverse, atgcttcgttgcacggatct. PCR products were digested with detection enzyme and analysed by DNA gel electrophoresis. Cleavage efficiency was calculated as: cleavage efficiency $= 1-[(1-\text{fraction cleaved})^{1/2}]$, where fraction cleaved = sum of cleaved band intensities/(sum of the cleaved and parental band intensities).

Data analysis. Offline data analysis for electrophysiology was performed using custom scripts in Igor Pro 6 (Wavemetrics) and MATLAB (MathWorks). Statistical analyses were performed using Prism 6 (GraphPad). Imaging data analyses were performed with ImageJ (NIH). For imaging results of FOS staining, ImageJ was used to convert original fluorescent images following a three-step method (grays, invert LUT, merge channels). All values are reported as mean ± s.e.m.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data and code availability. The original and/or analysed datasets generated during the current study, and the codes used to analyse them, are available from the corresponding author upon reasonable request.

31. Tong, Q., Ye, C.-P., Jones, J. E., Elmquist, J. K. & Lowell, B. B. Synaptic release of GABA by AGRP neurons is required for normal regulation of energy balance. *Nat. Neurosci.* **11**, 998–1000 (2008).
32. Parton, L. E. et al. Glucose sensing by POMC neurons regulates glucose homeostasis and is impaired in obesity. *Nature* **449**, 228–232 (2007).

33. Leiter, E. H. The NOD mouse: a model for analyzing the interplay between heredity and environment in development of autoimmune disease. *ILAR J.* **35**, 4–14 (1993).
34. Liu, T. et al. Fasting activation of AgRP neurons requires NMDA receptors and involves spinogenesis and increased excitatory tone. *Neuron* **73**, 511–522 (2012).
35. Ran, F. A. et al. Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
36. Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **42**, W401–W407 (2014).
37. Henry, F. E., Sugino, K., Tozer, A., Branco, T. & Sternson, S. M. Cell type-specific transcriptomics of hypothalamic energy-sensing neuron responses to weight-loss. *eLife* **4**, e09800 (2015).
38. Opland, D. et al. Loss of neurotensin receptor-1 disrupts the control of the mesolimbic dopamine system by leptin and promotes hedonic feeding and obesity. *Mol. Metab.* **2**, 423–434 (2013).
39. Kong, D. et al. A postsynaptic AMPK→p21-activated kinase pathway drives fasting-induced synaptic plasticity in AGRP neurons. *Neuron* **91**, 25–33 (2016).
40. Kong, D. et al. GABAergic RIP–Cre neurons in the arcuate nucleus selectively regulate energy expenditure. *Cell* **151**, 645–657 (2012).
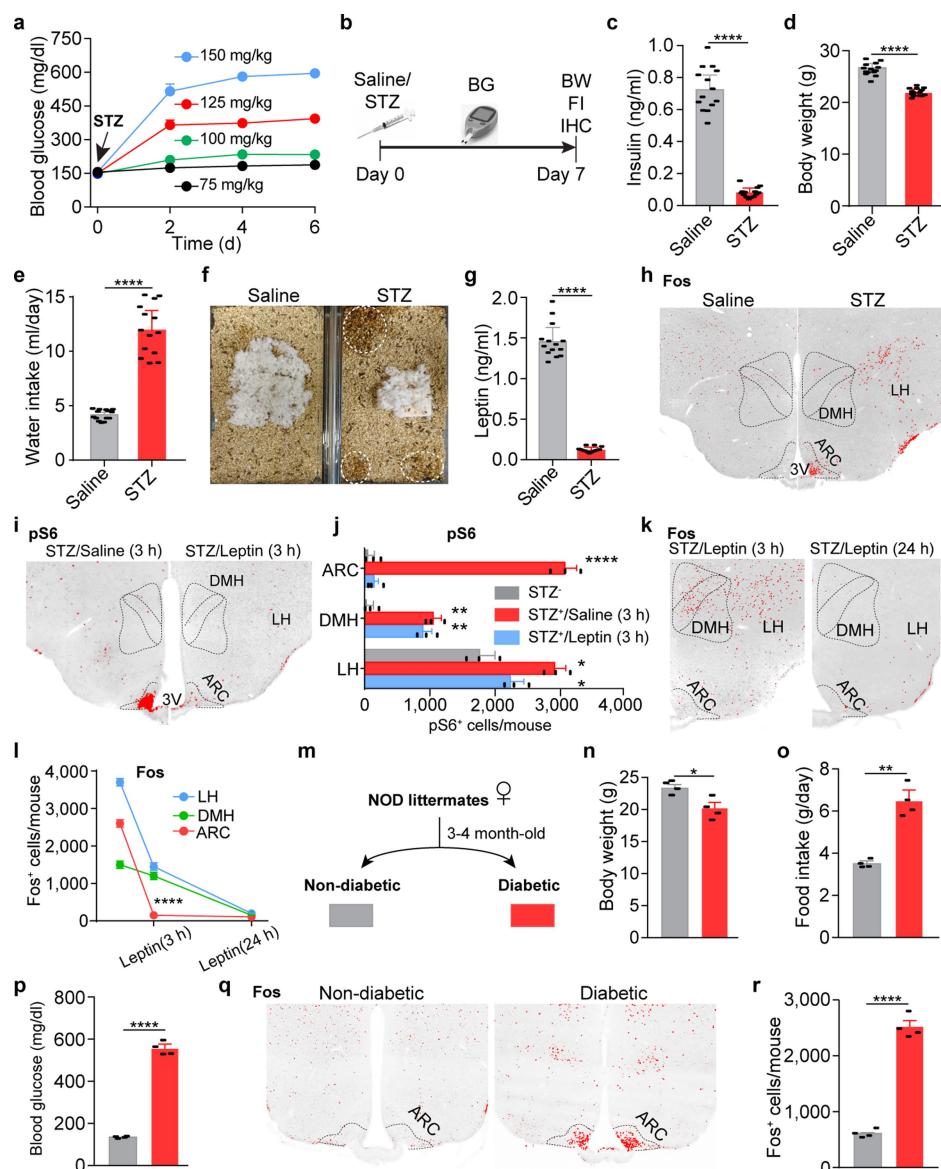
Extended Data Fig. 1 | Characterization of STZ-induced diabetic mice.
Additional analysis of alterations in neuronal activity in STZ-treated
animals and following leptin administration; similar activation of neurons
in the ARC was also observed in NOD diabetic mice. **a**, Development
of hyperglycaemia in C57BL/6 mice after one-time intraperitoneal
administration of STZ at different doses from 75 mg kg⁻¹ to 150 mg kg⁻¹
($n = 4$ mice per group). STZ treatments in subsequent experiments used a
dose of 125 mg kg⁻¹. **b**, Experimental protocol. **c–g**, Serum insulin levels (**c**),
body weight (**d**), daily water intake (**e**), representative cages in which a
saline or STZ-treated mouse were housed (**f**) and serum leptin levels (**g**)
after one week post-STZ-injection ($n = 14$ mice per group), suggesting
that treatment with 125 mg kg⁻¹ STZ effectively induces insulin deficiency
and diabetes, as well as emaciation, polydipsia, polyuria and leptin
deficiency. **h**, Representative sections of FOS immunostaining in the

hypothalamus of saline- or STZ-treated mice. **i, j**, Representative sections
and quantification of pS6 (Ser235,236) immunostaining in the mediobasal
hypothalamus of saline- or STZ-treated mice 3 h after the administration
of saline or leptin ($n = 3$ mice per group). **k, l**, Comparison of FOS
expression in the mediobasal hypothalamus of STZ-treated mice following
3-h or 24-h leptin treatment ($n = 3$ per group). **m**, Schematic diagram of
the development of the non-obese diabetic (NOD) mouse model.
n–r, Body weight (**n**), daily food intake (**o**), ad libitum-fed blood glucose
levels (**p**), and representative sections (**q**) and quantification (**r**) of FOS
immunostaining in the ARC of non-diabetic or diabetic NOD littermates
($n = 4$ mice per group). Data are mean ± s.e.m. and representative of three
independent experiments; *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$; Student's
two-tailed, unpaired t-test (**c–e, g, j, n–p, r**) or two-way ANOVA (**l**).
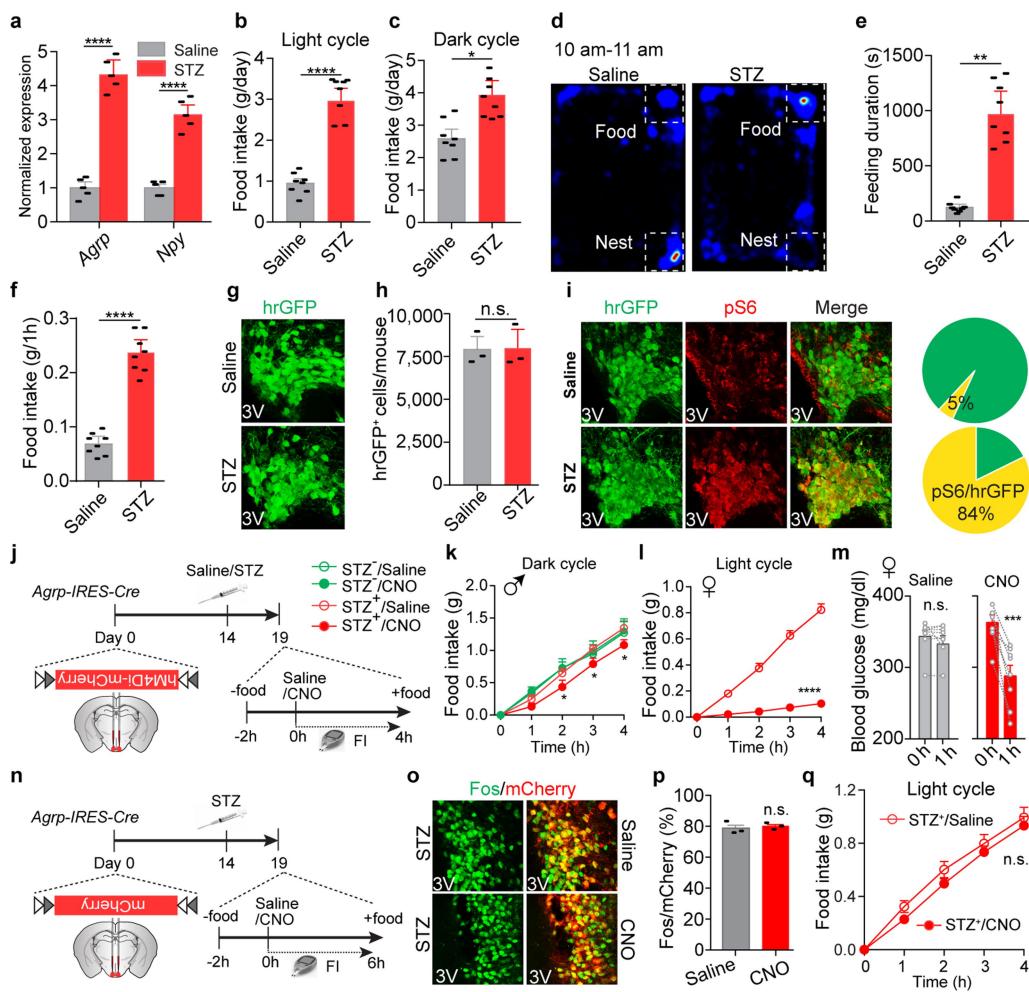
Extended Data Fig. 2 | Characterization of feeding behaviours in STZ-treated diabetic mice. Additional analyses of ectopic activation of AGRP neurons and its pathologic contributions to STZ-induced hyperphagia and hyperglycaemia. **a**, Quantitative PCR results showing that expression of *Agrp* and *Npy* is significantly upregulated in the mediobasal hypothalamus of STZ-treated animals, which is consistent with increased AGRP neuronal activity following STZ injection ($n = 5$ mice per group). **b–f**, Food intake during light cycle (07:00 to 19:00) (**b**) and dark cycle (19:00 to 07:00) (**c**), representative 1-h heat map (10:00 to 11:00) showing per cent occupancy time in food zone (upper right corner) and nesting zone (lower right corner) (**d**), feeding duration (**e**), and 1-h food intake (**f**) in saline- or STZ-treated C57BL/6 mice ($n = 8$ mice per group). **g, h**, Representative sections and quantification of hrGFP immunostaining in the ARC of saline- or STZ-treated *Npy-hrGFP* transgenic mice, suggesting that STZ treatment does not induce obvious loss of AGRP neurons ($n = 3$ mice per group). **i**, Representative sections and quantification of hrGFP and pS6 co-immunostaining in the ARC of saline- or STZ-treated *Npy-hrGFP* transgenic mice ($n = 4$ mice per group). **j**, Schematic of chemogenetic inhibition of AGRP neurons in virus-transduced *Agrp-IRES-cre* mice.

k, Four-hour food-intake measurement during dark cycle (20:00 to 00:00) following the administration of saline or CNO ($n = 6$ mice per group). **l, m**, Four-hour food-intake assay (10:00 to 14:00) (**l**) and blood glucose measurement (**m**, without food in the cage) in saline- or CNO-treated female *Agrp-IRES-cre* littermates that had been injected with AAV pSyn-FLEX-hM4Di-mCherry virus into the ARC, ($n = 8$ mice per group). **n–q**, Schematic of experiments to assess effects of CNO on *Agrp-IRES-cre* mice injected with Cre-dependent AAV-FLEX-mCherry virus in the ARC (**n**). Representative brain sections (**o**) and quantification (**p**, $n = 3$ mice per group) of mCherry and FOS co-immunostaining, and food intake assay (**q**; 10:00 to 14:00, $n = 8$ mice per group) in STZ-treated mice following intraperitoneal injection of saline or CNO, demonstrating that CNO administration without hM4Di expression in AGRP neurons has no effect and the changes observed in Fig. 1g–i are caused by the chemogenetic inhibition of AGRP neurons. Data are mean ± s.e.m. and representative of three independent experiments; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired *t*-test (**a–c, e, f, h, p**), paired *t*-test (**m**) or two-way ANOVA analysis (**l, q**) with Šidák post hoc test (**k**).
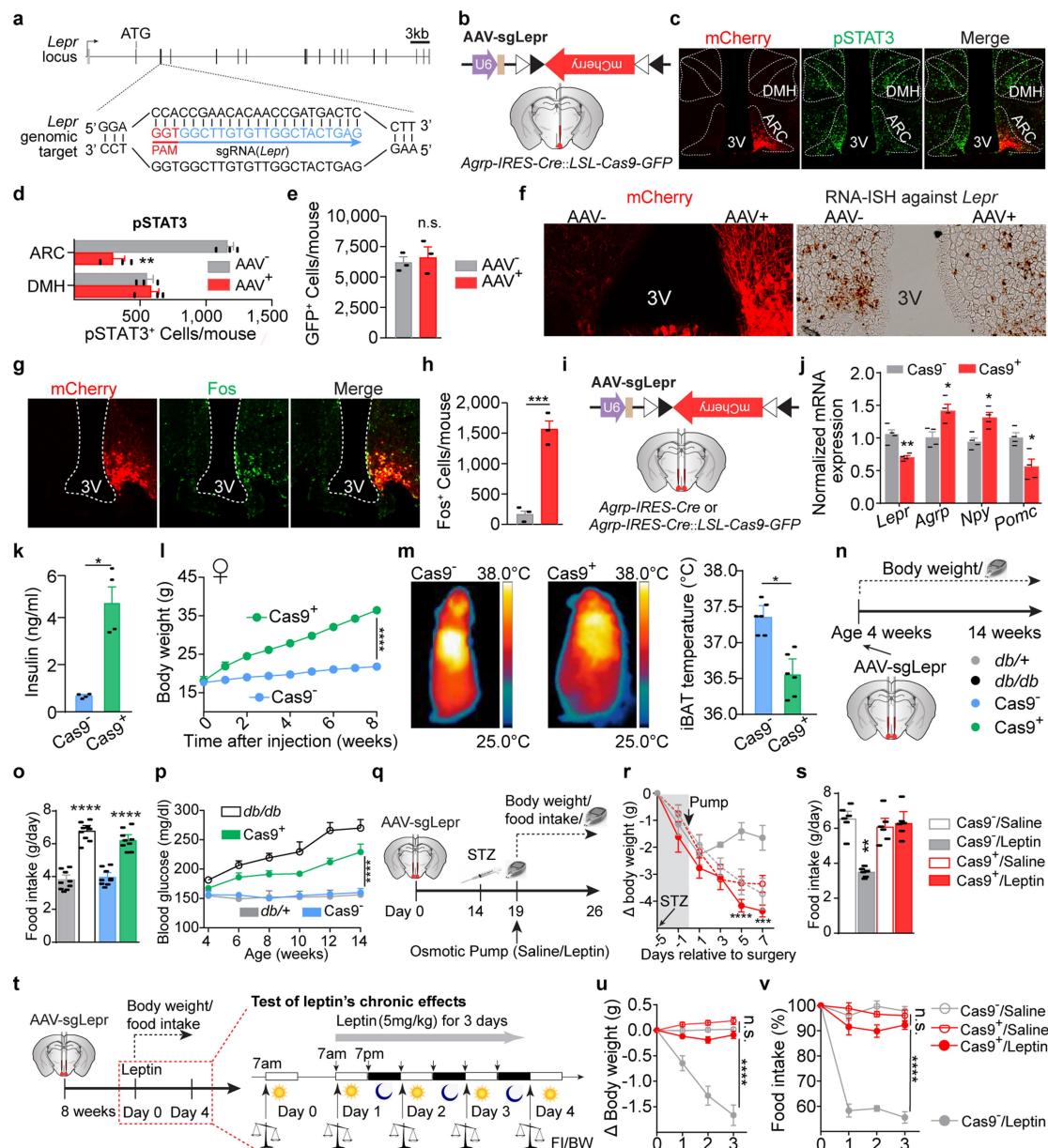
Extended Data Fig. 3 | Additional information and analyses for CRISPR-mediated disruption of LEPRs in AGRP neurons. a, sgRNA design for targeting the mouse *Lepr* genomic locus. **b–h**, AAV pU6-sgRNA[Lepr]::pEF1α-FLEX-mCherry (AAV-sgLepr) was injected unilaterally into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* mice (**b**). Representative images (**c**) and quantification (**d**) of mCherry and leptin-induced pSTAT3 (Tyr705) immunostaining in the hypothalamus. Cell counting of GFP[+] cells from the ARC suggests that deletion of *Lepr* in AGRP neurons does not induce cell death (**e**). Representative images of mCherry immunostaining (left) and ISH against *Lepr* mRNA (right) indicating efficient deletion of *Lepr* (**f**). Representative images (**g**) and quantification (**h**) of mCherry and FOS co-immunostaining in the ARC of ad libitum-fed mice, indicating disinhibition of AGRP neurons (*n* = 3 mice per group). **i, j**, AAV pU6-sgRNA[Lepr]::pEF1α-FLEX-mCherry (AAV-sgLepr) was injected bilaterally into the ARC of *Agrp-IRES-cre::LSL-Cas9-GFP* or *Agrp-IRES-cre* mice (**i**); reverse transcription with qPCR showing *Lepr*, *Agrp*, *Npy* and *Pomc* mRNA expression in the ventromedial hypothalamus of fed mice (**j**, *n* = 4 mice per group). **k**, Serum insulin levels in ad libitum-fed *Agrp-IRES-cre* (Cas9[−]) and *Agrp-IRES-cre::LSL-Cas9-GFP* mice (Cas9[+]) bilaterally injected with AAV-sgLepr virus (*n* = 4 mice per group). **l**, CRISPR-mediated deletion of *Lepr* in AGRP neurons also induces

severe obesity in female mice (*n* = 6 mice per group). **m**, Representative near-infrared thermal images and quantification of iBAT temperature in virus-transduced, ad libitum-fed *Agrp-IRES-cre* (Cas9[−]) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9[+]) male littermates. **n**, Assay design to compare mice with global *Lepr* mutations and mice with specific deletion of *Lepr* in AGRP neurons. **o, p**, Weekly blood glucose measurement and daily food intake at eight weeks of age of ad libitum-fed, virus-transduced *Lepr[db/+]*, *Lepr[db/db]*, *Agrp-IRES-cre* (Cas9[−]), and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9[+]) mice (*n* = 9 mice per group). **q–s**, Experimental design (**q**), changes in body weight (**r**) and daily food intake (**s**) five days post pump surgery in STZ-treated, virus-transduced *Agrp-IRES-cre* (Cas9[−]) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9[+]) mice following chronic administration of saline or leptin with osmotic pump (*n* = 6 mice per group). **t–v**, Experimental design (**t**), changes of body weight (**u**) and daily food intake (**v**) in non-STZ treated, virus-transduced *Agrp-IRES-cre* (Cas9[−]) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9[+]) mice following three-day intraperitoneal injection of leptin (*n* = 7 mice per group). Data are mean ± s.e.m. and representative of three independent experiments; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired *t*-test (**d, e, h, j, k, m, o, s**) or two-way ANOVA (**l, p, u, v**) with Šidák post hoc test (**r**).
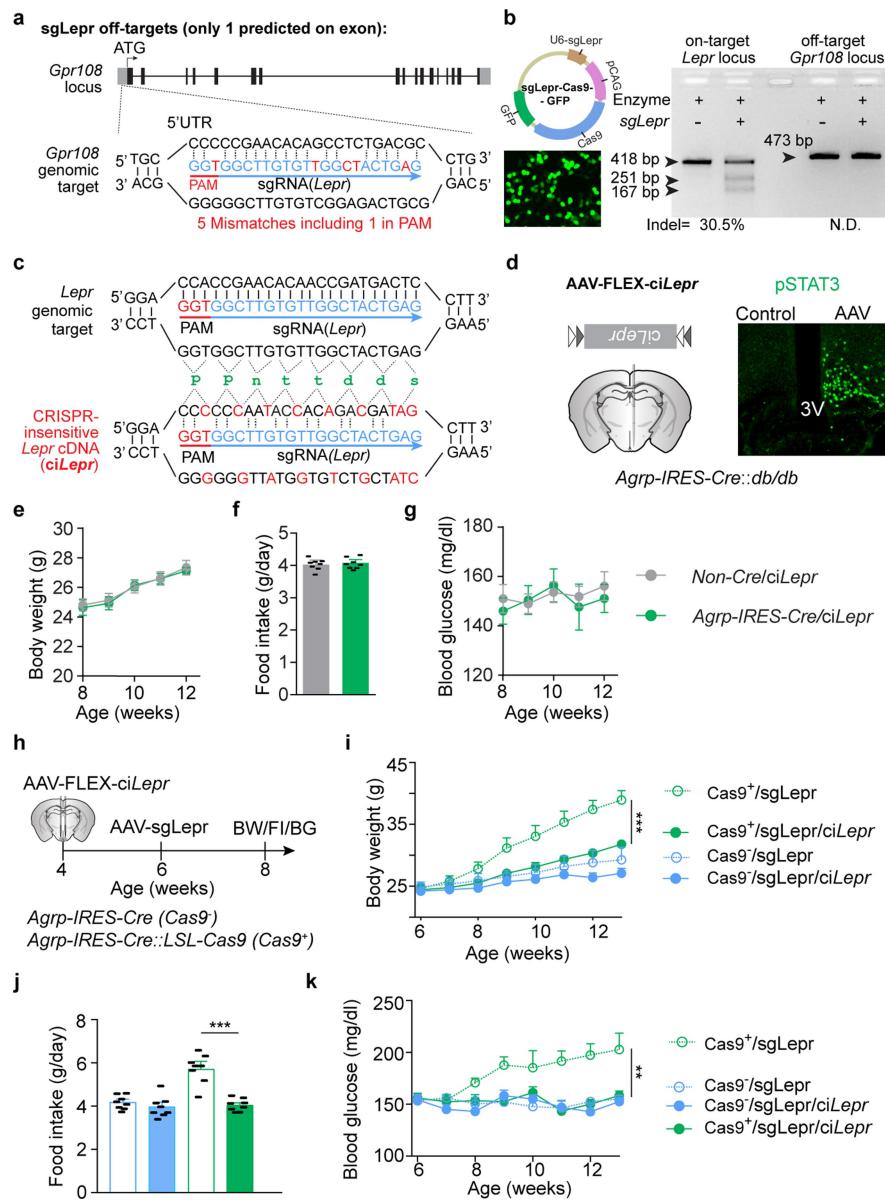
Extended Data Fig. 4 | Analysis of off-target effects of CRISPR-mediated genome editing. Expression of CRISPR-insensitive leptin receptors in AGRP neurons occludes alterations in body weight, food intake or blood glucose levels caused by CRISPR-mediated disruption of *Lepr*. **a**, Predicted off-target site of sgLepr on the first exon of *Gpr108* locus. **b**, Plasmid for co-expression of sgLepr and Cas9-GFP proteins (top left); GFP immunofluorescence following transfection of the plasmid into mouse N2a cells (bottom left); on-target and off-target indel detection, demonstrating that sgLepr effectively induces mutations in the *Lepr* locus but not in the *Gpr108* locus (right). **c**, sgRNA targeting the mouse *Lepr* locus (upper) and CRISPR-insensitive *Lepr* cDNA encoding the long-form leptin receptors (ci*Lepr*, lower) with indicated silent mutations to prevent binding of sgRNA. **d**, Cre-dependent AAV pEF1α-FLEX-ci*Lepr* (AAV-FLEX-ci*Lepr*) injected unilaterally into the ARC of *Agrp-IRES-cre::Lepr^db/db^* mice (left) and representative images of leptin-induced

pSTAT3 (Tyr705) immunostaining (right), indicating that ci*Lepr* is fully responsive to leptin stimulation. **e–g**, Body weight (**e**), food intake (**f**) and blood glucose (**g**) of non-*cre* and *Agrp-IRES-cre* littermates following bilateral injection of AAV-FLEX-ci*Lepr* into the ARC ($n = 8$ mice per group), suggesting that AAV-mediated expression of ci*Lepr* in AGRP neurons produces no obvious effects on energy or glucose balance. **h–k**, Experimental protocol (**h**), body weight (**i**), daily food intake (**j**) and blood glucose measurements (**k**) in virus-transduced ad libitum-fed littermates ($n = 8$ mice per group). Since expression of ci*Lepr* in AGRP neurons prevented body weight gain, increased food intake and hyperglycaemia induced by CRISPR-mediated deletion, potential contributions from off-site mutagenesis are excluded. Data are mean ± s.e.m. and representative of three independent experiments; **$P < 0.01$, ***$P < 0.001$; Student's two-tailed, unpaired t-test (**f**, **j**) or two-way ANOVA (**e**, **g**, **i**, **k**).
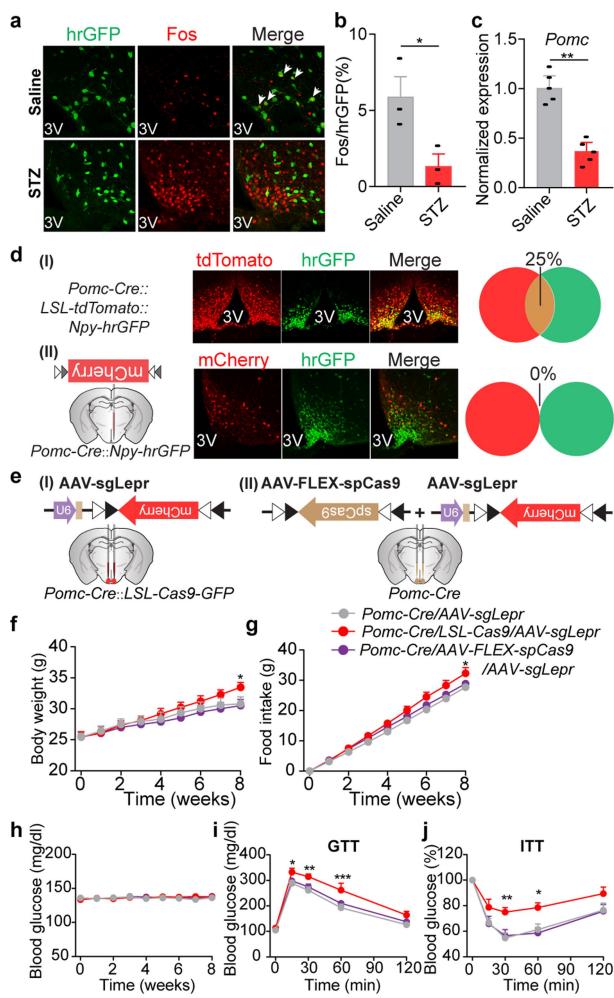
Extended Data Fig. 5 | CRISPR-mediated disruption of leptin receptors in hypothalamic POMC neurons does not alter energy or blood glucose balance. a, b, Representative sections and quantification of hrGFP and FOS co-immunostaining in the ARC of saline- or STZ-treated *Pomc-hrGFP* transgenic mice (*n* = 3 mice per group). Reduced FOS expression in POMC neurons was observed following STZ treatment. **c,** Results of reverse transcription with qPCR, showing that *Pomc* mRNA levels are significantly reduced in the mediobasal hypothalamus of STZ-treated animals, which is consistent with inhibited POMC neuronal activity following STZ injection (*n* = 5 mice per group). **d,** Schematics, representative sections and quantification of tdTomato and hrGFP co-immunostaining in the ARC of *Pomc-cre::LSL-tdTomato::Npy-hrGFP* mice (top) and mCherry and hrGFP co-immunostaining in the ARC of *Pomc-cre::Npy-hrGFP* mice with Cre-dependent AAV pEF1α-FLEX-mCherry injected into the ARC (bottom). Co-expression was observed in *Pomc-cre::LSL-tdTomato::Npy-hrGFP* mice but not in virus-transduced *Pomc-cre::Npy-hrGFP* mice, suggesting that *Pomc-cre* ectopically induces Cre activity in AGRP neurons at an early developmental stage, consistent with previous findings. These data also demonstrate that Cre-dependent AAV injected into the ARC of *Pomc-cre* transgenic mice provides an efficient approach to specifically express genes of interest in POMC ARC neurons, without perturbing intermingled AGRP neurons (*n* = 3 mice per group). **e,** Two approaches to achieve CRISPR-mediated deletion of *Lepr* in POMC neurons. Left, AAV-sgLepr virus was bilaterally injected into the ARC of *Pomc-cre::LSL-Cas9-GFP* mice. Right, a viral mix of AAV-sgLepr and Cre-dependent AAV pMeCP2-FLEX-spCas9 (AAV-FLEX-spCas9, see Methods) was bilaterally injected into the ARC of *Pomc-cre* mice. **f–j,** Body weight (**f**), accumulated weekly food intake (**g**), ad libitum-fed-state blood glucose levels (**h**), glucose-tolerance test (**i**) and insulin-tolerance test (**j**) in the virus-transduced animals (*n* = 6 mice per group). Notably, single virus-transduced *Pomc-Cre::LSL-Cas9-GFP* mice, but not dual virus-transduced *Pomc-Cre* mice, exhibited mildly increased body weight and food intake, and slightly impaired glucose tolerance and insulin sensitivity. Since Cas9 protein would be expected to be expressed in some AGRP neurons in *Pomc-cre::LSL-Cas9-GFP* mice, the difference observed between the two approaches is likely to be explained by the ectopic Cre activity of *Pomc-cre* in AGRP neurons. The phenotypes observed in *Pomc-cre::LSL-Cas9-GFP* mice mimic those following genetic ablation of *Lepr* with a conventional Cre–loxP system. Inactivation of LEPR in POMC neurons of adult mice does not appear to affect energy balance or glucose homeostasis under the assayed conditions. Data are mean ± s.e.m. and representative of three independent experiments; *P < 0.05, **P < 0.01, ***P < 0.001; Student's two-tailed, unpaired *t*-test (**b**, **c**) or two-way ANOVA with Šidák post hoc test (**f–j**).
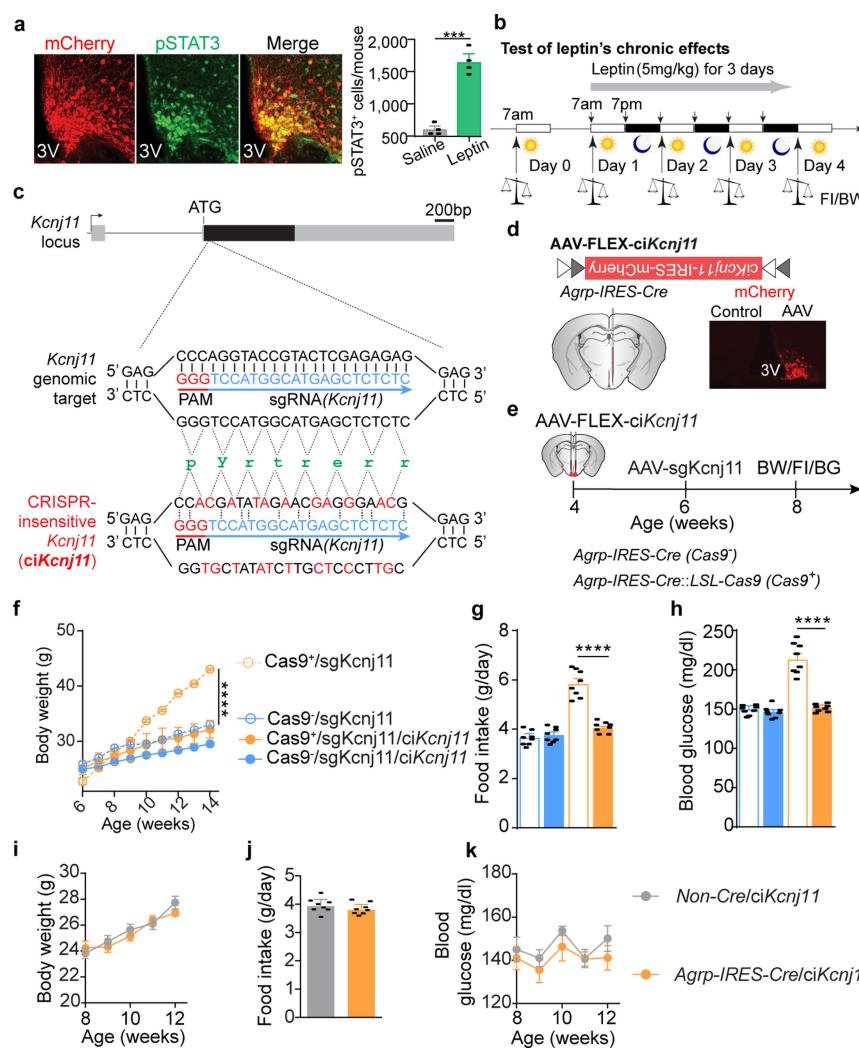
Extended Data Fig. 6 | Additional information on experimental design related to CRISPR-mediated deletion of *Kcnj11* in AGRP neurons. Expression of a CRISPR-insensitive Kir6.2 in AGRP neurons prevents alterations in body weight, food intake or blood glucose following CRISPR-mediated deletion. **a**, Representative images and quantification of mCherry and pSTAT3 co-immunostaining ($n = 4$ mice per group) following CRISPR-mediated disruption of *Kcnj11* in AGRP neurons. **b**, Experimental protocol to test chronic effects of leptin on regulation of body weight and food intake. **c**, sgRNA design targeting the mouse *Kcnj11* locus and CRISPR-insensitive *Kcnj11* cDNA encoding the mouse K_{ATP} channel subunit Kir6.2 (ci*Kcnj11*) with indicated silent mutations to prevent binding of sgRNA. **d**, Schematic of Cre-dependent AAV pEF1α-FLEX-ci*Kcnj11*-IRES-mCherry (AAV-FLEX-ci*Kcnj11*) injected unilaterally into the ARC of *Agrp-IRES-cre::Lepr*$^{db/db}$ mice and representative image of mCherry immunostaining. **e–h**, Experimental protocol (**e**), body weight (**f**), daily food intake (**g**) and blood glucose measurements (**h**) in virus-transduced ad libitum-fed littermates ($n = 8$ mice per group). Expression of ci*Kcnj11* in AGRP neurons completely prevented body weight gain, increased food intake and hyperglycaemia induced by CRISPR-mediated deletion, suggesting that contributions from off-site mutagenesis are excluded. **i–k**, Body weight (**i**), food intake (**j**), and blood glucose (**k**) of non-*cre* and *Agrp-IRES-cre* littermates following bilateral injection of AAV-FLEX-ci*Kcnj11* into the ARC ($n = 8$ mice per group), suggesting that AAV-mediated expression of ci*Kcnj11* in AGRP neurons produces no obvious effects on energy or glucose balances. Data are mean ± s.e.m. and representative of three independent experiments; ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired t-test (**a**, **g**, **h**, **j**) or two-way ANOVA (**f**, **i**, **k**).
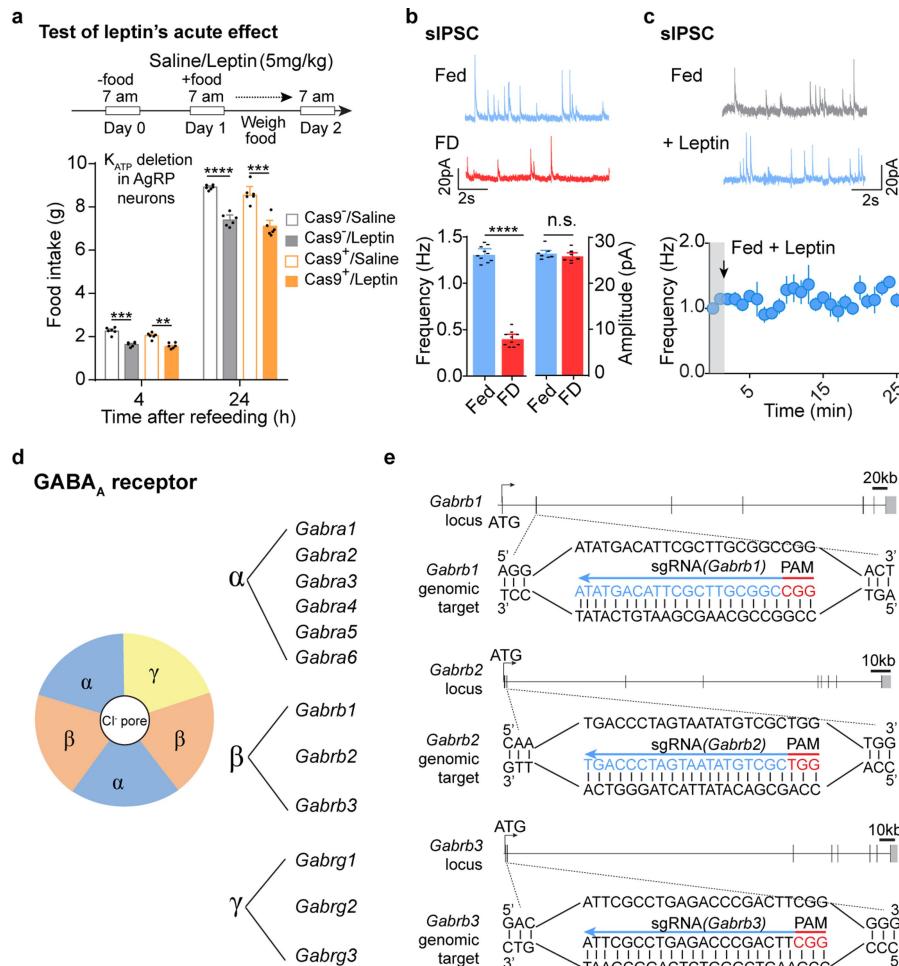
Extended Data Fig. 7 | Additional characterization of inhibition of fasting-induced hunger by leptin. Fasting- and leptin-induced alterations in GABAergic neurotransmission on AGRP neurons. **a**, Assays to test acute effects of leptin in suppressing 24-h fasting-induced hunger in *Agrp-IRES-cre* (Cas9−) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9+) mice, following bilateral injection of AAV-sgK$_{ATP}$ into the ARC. These data indicate that K$_{ATP}$ channels in AGRP neurons are not required for acute inhibition of hunger by leptin following fasting. $n = 6$ mice per group. **b**, Representative traces of spontaneous inhibitory postsynaptic currents (sIPSCs) on AGRP neurons and quantification of their frequency and amplitude in ad libitum-fed or 24-h-fasted *Npy-hrGFP* transgenic mice. No difference was observed in sIPSC amplitude. $n = 10$ cells from three mice per group. **c**, Representative traces and quantification of sIPSCs on AGRP neurons before or 15 min post-incubation with leptin in ad libitum-fed *Npy-hrGFP* transgenic mice, suggesting that leptin does not further regulate GABAergic neurotransmission on AGRP neurons in fed animals. $n = 10$ cells from three mice per group. **d**, Schematic of the mouse GABA$_A$ receptor complex and the genes encoding its subunits. **e**, Genomic structures and design of sgRNAs targeting the mouse *Gabrb1–Gabrb3* loci encoding GABA$_A$ receptor subunits β1 to β3, respectively. Data are mean ± s.e.m. and representative of three independent experiments; **P < 0.01, ***P < 0.001, ****P < 0.0001; Student's two-tailed, unpaired *t*-test (**a**, **b**).
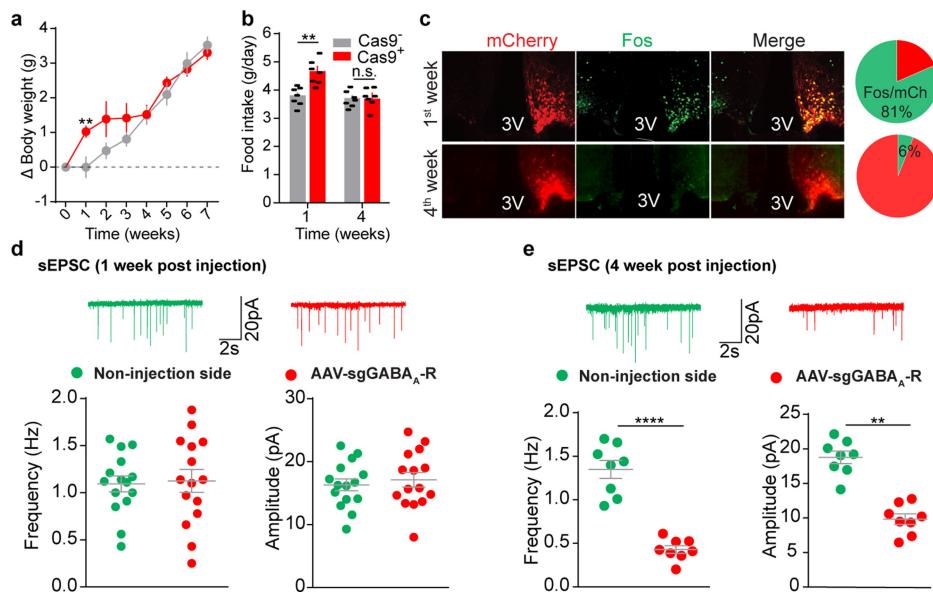
Extended Data Fig. 8 | Dynamic changes of neuronal activity and synaptic neurotransmission following CRISPR-mediated deletion of GABA$_A$ receptors in AGRP neurons. a, b, Body weight (**a**) and daily food intake (**b**) of *Agrp-IRES-cre* (Cas9$^-$) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9$^+$) mice following bilateral injection of AAV-sgGABA$_A$-R into the ARC ($n = 7$ mice per group). **c–f,** Representative sections and quantification of mCherry and FOS co-immunostaining (**c**, $n = 3$ mice per group), representative traces and quantification of spontaneous excitatory postsynaptic currents (sEPSCs) (**d**, $n = 15$ neurons from three mice; **e**, $n = 8$ neurons from two mice) on AGRP neurons of *Agrp-IRES-cre::LSL-Cas9-GFP* mice after one-week or four-week unilateral injection of AAV-sgGABA$_A$-R into the ARC. Reduced frequency and amplitude of sEPSCs were observed in AGRP neurons at a later stage following GABA$_A$ receptor deletion, together with concurrent elimination of the observed disinhibition of these neurons (**c**), body weight gain (**a**) and hyperphagia (**b**), suggesting that excitatory and inhibitory afferents cooperate dynamically to modulate AGRP neuronal activities. In addition, the compensatory reduction in sEPSCs in virus-transduced neurons but not in the neurons from the contralateral side also indicates a cell-autonomous mechanism. Data are mean ± s.e.m. and representative of three independent experiments; **P < 0.01, ****P < 0.0001; Student's two-tailed, unpaired *t*-test (**b, d, e**) or two-way ANOVA with Šidák post hoc test (**a**).
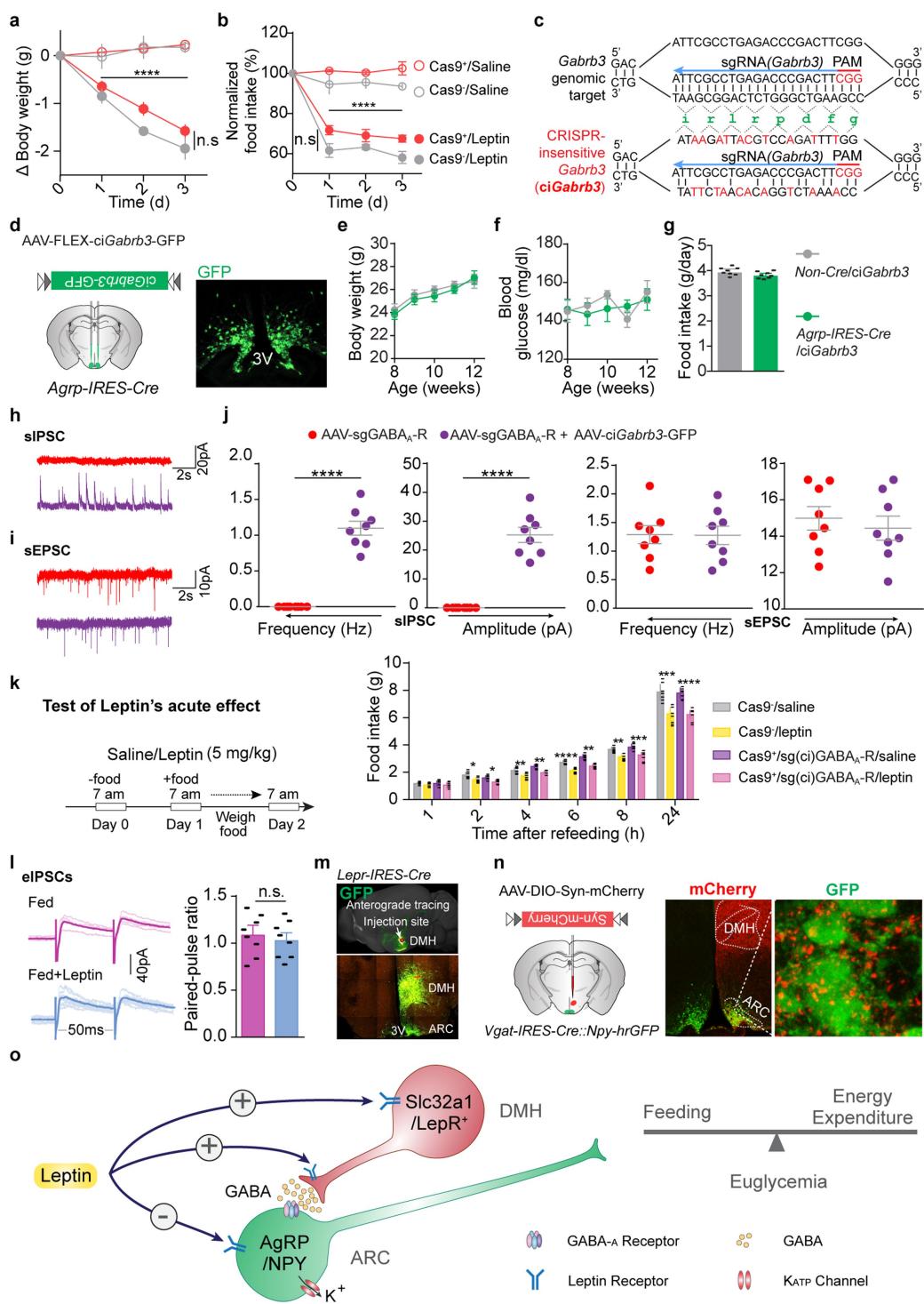
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Additional characterization of leptin regulation of energy balance and GABAergic neurotransmission on AGRP neurons following CRISPR-mediated deletion or expression of CRISPR-insensitive GABA_A receptors. a, b, Body weight and daily food intake changes during three-consecutive-day treatment with saline or leptin in *Agrp-IRES-cre* (Cas9⁻) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9⁺) mice following bilateral injection of AAV-sgGABA_A-R into the ARC ($n = 7$ mice per group), suggesting that GABAergic neurotransmission on AGRP neurons is not required for chronic effects of leptin on regulation of body weight or food intake. **c,** sgRNA design targeting the mouse *Gabrb3* locus and CRISPR-insensitive *Gabrb3* cDNA encoding the mouse β3 subunit of GABA_A receptors (ci*Gabrb3*) with indicated silent mutations to prevent binding of sgRNA. **d,** Schematic of Cre-dependent AAV pEF1α-FLEX-ci*Gabrb3*-GFP (AAV-FLEX-ci*Gabrb3*) and representative image of GFP immunostaining following bilateral injection of the virus into the ARC of *Agrp-IRES-cre* mice. **e–g,** Body weight (**e**), food intake (**f**) and blood glucose (**g**) of non-*cre* and *Agrp-IRES-cre* littermates following bilateral injection of AAV-FLEX-ci*Gabrb3* into the ARC ($n = 8$ mice per group), suggesting that AAV-mediated expression of ci*Gabrb3* in AGRP neurons produces no obvious effects on energy or glucose balances. **h–j,** Representative traces and quantification of sIPSCs and sEPSCs of AGRP neurons in *Agrp-IRES-cre::LSL-Cas9-GFP* mice, following bilateral injection of AAV-sgGABA_A-R (red) or a mix of AAV-sgGABA_A-R and AAV-FLEX-ci*Gabrb3* (purple) into the ARC ($n = 8$ neurons from three mice per group). **k,** Assays to test acute effects of leptin in suppressing 24-h-fasting-induced hunger in *Agrp-IRES-cre* (Cas9⁻) and *Agrp-IRES-cre::LSL-Cas9-GFP* (Cas9⁺) mice, following bilateral injection of a mix of AAV-sgGABA_A-R and AAV-FLEX-ci*Gabrb3* into the ARC. These data indicate that expression of CRISPR-insensitive β3 subunit of GABA_A receptor in AGRP neurons following deletion of GABA_A receptor subunits β1–β3 restores acute inhibition of hunger by leptin following fasting ($n = 7$ mice per group). Functional contributions from CRISPR-mediated off-site mutagenesis of the triple deletion can therefore be excluded. **l,** Representative traces (left) and quantification (right) of electrically evoked IPSC (eIPSC) paired-pulse ratios in AGRP neurons of ad libitum-fed *Npy-hrGFP* mice before or after leptin incubation ($n = 8$ neurons from three mice per group). **m,** Representative anterograde tracing image (image from Allen Institute, experiment #113314337-DMH) showing intensive GFP-labelled projections in the ARC following unilateral viral injections in DMH of *Lepr-IRES-cre* mice. **n,** Schematic of unilateral injection of Cre-dependent AAV pEF1α-FLEX-synaptophysin-mCherry into the ventral DMH (vDMH) of *Vgat-IRES-cre::Npy-hrGFP* mice (left) and mCherry and hrGFP co-immunostaining (middle) to enable anterograde tracing of the projections of vDMH GABAergic neurons (vGAT^vDMH neurons). vGAT^vDMH neurons predominantly project to the ARC and their axon terminal puncta are intensively stained on the soma of AGRP neurons (right). **o,** Schematic summarizing leptin action on AGRP neurons and on GABAergic neurons in the DMH. Data are mean ± s.e.m. and representative of three independent experiments; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; Student's two-tailed, unpaired *t*-test (**g, j, k, l**) or two-way ANOVA (**a, b, e, f**).

Extended Data Table 1 | Summary of FOS expression in saline- or STZ-treated C57BL/6 mice

		Saline	STZ
Cortex	Frontal association cortex	+	-
	Dorsal peduncular cortex	+	-
	Infralimbic cortex	+	+++
	Prelimbic cortex	+	++++
	Cingulate cortex	++	+++
	Orbital cortex, lateral	++	+++
	Orbital cortex, ventral/medial	++	++++
	Orbital cortex, dorsal	+	+++
	Motor cortex, primary	++	++
	Motor cortex, secondary	++	++
	Agranular insular cortex	++	+
	Dysgranular/granular insular cortex	+	+
	Somatosensory cortex	++	++
	Auditory cortex	+	++
	Visual cortex	++	++++
	Entorhinal cortex	+	++
	Retrosplenial cortex	++	+
	Perirhinal cortex	+	+
	Ectorhinal cortex	+	+
	Parietal cortex	++	+++
Olfactory areas	Olfactory bulb	+++	+++++
	Anterior olfactory nucleus	++	++++
	Piriform cortex	+++	+++++
	Tenia tecta	-	+
Hippocampus/Septum	Hippocampus (CA1, CA2, CA3)	+	++++
	Dorsal Subiculum	-	-
	Hippocampus (dentate gyrus)	+	+
	Septum	-	-
	Diagonal band of Broca	++	++
Claustrum/endopiriform	Claustrum	++	-
	Endopiriform nucleus	+	-
Striatum	Olfactory tubercle	++	-
	N.Accumbens, Shell & Core	-	-
	Dorsal striatum	-	-
Pallidum	Ventral pallidum	-	-
	Globus pallidus	+	-
	Entopeduncular nucleus	-	-
	IPAC	+	+
	BNST	+	-
	Sublenticular extended amygdala	-	-
Amygdala	Central nucleus	-	+++
	Anterior amygdaloid area	-	-
	Medial nucleus	-	-
	Cortical amygdaloid transition zone	-	-
	Basolateral nucleus	++	+++
	Basomedial nucleus	+	+
Hypothalamus	Median preoptic area	+	+++++
	Medial preoptic area	+	++++
	Lateral preoptic area	+	+++
	Suprachiasmatic nucleus	++	+++++
	Arcuate nucleus	+	+++++
	Magnocellular preoptic area	-	+++
	Anterior hypothalamic area	+	+++
	Paraventricular hypothalamic nucleus	+	+++++
	Ventromedial hypothalamic nucleus	++	+
	Lateral hypothalamic area	+	+++
	Posterior hypothalamic area	++	+++
	Dorsal hypothalamic area	+	++++
	Tuber cinereum area	-	+++
	Perifornical nucleus	-	++
	Mammillary nucleus	-	+
	Premammillary nucleus	+	++
	Supramammillary nucleus	+	+++
	Subthalamic nucleus	+	+++
	Zona incerta	-	-
	Supraoptic nucleus	+	+++++
	Field of Forel, Prerubral field	-	-
	Retrochiasmatic nucleus	+	+
Thalamus/Epithalamus	Lateral habenula	-	++
	Medial habenula	-	-
	Parafascicular thalamic nucleus	-	+
	Paraventricular thalamic nucleus	+++	++++
	Lateral geniculate nucleus	+	++
	Medial geniculate nucleus	+	++
	Mediodorsal thalamic nucleus	-	-
	Central medial thalamic nucleus	-	+
	Ventromedial thalamic nucleus	-	++
	Dorsolateral thalamic nucleus	-	-
Midbrain	Olivary pretectal nucleus	-	++
	Nucleus of posterior commissure	-	+
	Nucleus of the optic tract	-	+
	Anterior pretectal nucleus	-	-
	Interpeduncular nucleus	-	+
	Substantia nigra pars reticulata	-	-
	Substantia nigra pars compacta	-	-
	Ventral tegmental area	-	+
	Retrorubral field	-	-
	Reticular formation	+	+
	Periaqueductal gray	+	++
	Superior colliculus	+	+
	Inferior colliculus	++	++
	Red nucleus	-	-
	Anterotegmental nucleus	-	-
	Median raphe nucleus	-	-
	Dorsal raphe nucleus	-	-
	Pontine reticular nucleus	-	-
	Pedunculopontine tegmental nucleus	-	-
	Laterodorsal tegmental nucleus	-	-
	Subpeduncular tegmental nucleus	-	-
	Retrorubral nucleus	-	-
	Cuneiform nucleus	-	-
	Parageminal nucleus	-	-
	Dorsal tegmental nucleus	-	-
	Pontine raphe nucleus	+	-
Hindbrain	Locus coeruleus	+	++++
	Superior olive	-	-
	Parabrachial nucleus	+	++
	Raphe magnus nucleus	-	-
	Raphe pallidus nucleus	-	++
	Solitary nucleus	-	++
	Paragigantocellular reticular nucleus	-	++

"+" or "-" indicates the intensity of Fos immune-positive neurons in each brain region. "+" indicates the lowest level of expression and "+++++" indicates the highest. "-" indicates the absence of detected expression.

LETTER

Pluripotency factors functionally premark cell–type–restricted enhancers in ES cells

Hong Sook Kim[1]*, Yuliang Tan[1,3], Wubin Ma[1,3], Daria Merkurjev[1,3], Eugin Destici[1,3], Qi Ma[1], Tom Suter[1], Kenneth Ohgi[1], Meyer Friedman[1], Dorota Skowronska–Krawczyk[1,2] & Michael G. Rosenfeld[1]*

Enhancers for embryonic stem (ES) cell-expressed genes and lineage-determining factors are characterized by conventional marks of enhancer activation in ES cells[1–3], but it remains unclear whether enhancers destined to regulate cell-type-restricted transcription units might also have distinct signatures in ES cells. Here we show that cell-type-restricted enhancers are 'premarked' and activated as transcription units by the binding of one or two ES cell transcription factors, although they do not exhibit traditional enhancer epigenetic marks in ES cells, thus uncovering the initial temporal origins of cell-type-restricted enhancers. This premarking is required for future cell-type-restricted enhancer activity in the differentiated cells, with the strength of the ES cell signature being functionally important for the subsequent robustness of cell-type-restricted enhancer activation. We have experimentally validated this model in macrophage-restricted enhancers and neural precursor cell (NPC)-restricted enhancers using ES cell-derived macrophages or NPCs, edited to contain specific ES cell transcription factor motif deletions. DNA hydroxyl-methylation of enhancers in ES cells, determined by ES cell transcription factors, may serve as a potential molecular memory for subsequent enhancer activation in mature macrophages. These findings suggest that the massive repertoire of cell-type-restricted enhancers are essentially hierarchically and obligatorily premarked by binding of a defining ES cell transcription factor in ES cells, dictating the robustness of enhancer activation in mature cells.

Enhancers function as critical regulatory elements that integrate genomic information for cell fate transition and cell-specific gene regulation[1–3]. We hypothesized that cell-type-restricted enhancers might be premarked in ES cells. To begin to explore this question, we selected macrophage enhancers, because the sequential events leading to macrophage differentiation and regulation by inflammatory signals are relatively well understood[4–6]. We first examined the full repertoire of macrophage-restricted enhancers in mouse ES cells and found that the majority (18,405 out of 22,684 enhancers) lacked H3K4me1 (monomethylation of histone H3 at lysine 4), H3K4me2 (dimethylation of histone H3 at lysine 4), H3K27Ac (acetylation of histone H3 at lysine 27) and H3K27me3 (trimethylation of histone H3 at lysine 27) marks (referred to as 'unmarked' in Extended Data Fig. 1a), although about 4,000 enhancers that were active in both macrophages and ES cells, including 'housekeeping genes', did exhibit H3K4me2 and H3K27Ac, generally within 200 kb of coding target genes (Fig. 1a, Extended Data Fig. 1a). Finally, a small number of enhancers (214) carried H3K27me3, which marks 'poised' enhancers (Fig. 1a, Extended Data Fig. 1a). We carried out an assay for transposase-accessible chromatin using sequencing (ATAC-seq) to further understand the enhancer features in ES cells. The 18,405 unmarked macrophage enhancers were in an open configuration compared to random regions (Fig. 1b), consistent with published DNase hypersensitivity analyses in ES cells (Fig. 1b), but were not as robustly accessible as enhancers that are active in ES cells (Extended Data Fig. 1b). To identify transcription factors that might bind to macrophage enhancers in open chromatin configurations, we profiled the distribution of several of the most important ES cell transcription factors—ESRRB, NANOG, OCT4 and SOX2 (ENOS)—in a $-1\,\mathrm{kb}/+1\,\mathrm{kb}$ window, ensuring that we were exclusively analysing macrophage-restricted enhancers. Notably, 6,775 macrophage-restricted enhancers showed binding of ENOS (Fig. 1c). We established the specificity of ENOS binding in macrophage-restricted enhancers by comparing with random regions (Extended Data Fig. 1c), which revealed statistically significant binding of ESRRB (Extended Data Fig. 1d). Notably, about 80% of macrophage-restricted enhancers were bound by only one or at most two ES cell transcription factors, whereas ES cell-active enhancers were bound by all four ENOS factors (Fig. 1d, Extended Data Fig. 1e), as exemplified by genome browser images (Extended Data Fig. 1f). We also analysed 12 ES cell transcription factors from the published literature[7], and found that active ES cell-restricted enhancers were characteristically bound predominantly by about 4–8 of the 12 ES cell transcription factors evaluated (OCT4, SOX2, NANOG, ESRRB, SMAD1, E2F1, TCFCP2L1, ZFX, STAT3, KLF4, C-MYC and N-MYC), consistent with their reported cooperative binding[7,8], while the majority of the active macrophage-restricted enhancers exhibited binding of only one or two of these factors (Fig. 1e).

To determine whether cell-type-restricted enhancers in other cell types also exhibit similar pre-marking, we examined cell-type-restricted enhancers from heart, kidney and N2A neuronal cells in a mouse model. These enhancers in ES cells again predominantly exhibited binding of a single ENOS factor and chromatin openness (Extended Data Fig. 2a–d).

Given the established role of the cohesin complex in chromatin architecture and gene regulation[9–11], we investigated whether this complex has a role in premarked enhancers. Consistent with a previous report[11], cohesin was colocalized with ENOS-bound regions (Fig. 2e). Therefore, we next investigated whether premarked enhancers could interact with other genomic regions. We performed circular chromatin conformation capture followed by deep sequencing (4C–seq) on a macrophage enhancer located 5′ of Il1a[12]. This premarked enhancer interacted specifically with other genomic regions in ES cells, including an upstream CTCF-bound site, but not with the cognate promoter. In macrophages treated with the bacterial endotoxin Kdo2-lipid A (KLA), this enhancer interacts robustly with the cognate promoter of the target coding gene (Extended Data Fig. 3a). An enhancer located 5′ of Tnfaip3 was similarly found to interact specifically with other genomic regions in ES cells, but not with its cognate promoter (Extended Data Fig. 3b). In addition, we performed RNA sequencing (RNA-seq) and examined the proximal macrophage-expressed genes of 6,775 macrophage-restricted enhancers. We found 634 differentially expressed genes (fold-change $\geq$ 4), which corresponded to functional categories relevant to macrophages but not ES cells (Extended Data Fig. 3c, d). These data, together with the 4C–seq results, suggest that premarked macrophage enhancers are not functional, despite their binding of ES cell transcription factors and open chromatin configurations in ES cells.

[1]Howard Hughes Medical Institute, Department and School of Medicine, University of California San Diego, La Jolla, CA, USA. [2]Present address: Shiley Eye Institute, Department of Ophthalmology, University of California San Diego, La Jolla, CA, USA. [3]These authors contributed equally: Yuliang Tan, Wubin Ma, Daria Merkurjev, Eugin Destici. *e-mail: hsk020@ucsd.edu; mrosenfeld@ucsd.edu
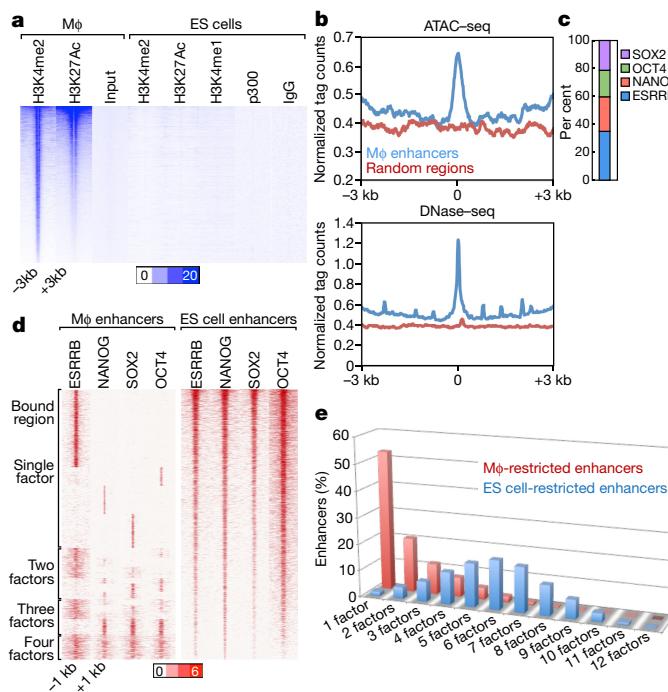
Fig. 1 | Features of macrophage enhancers in ES cells. a, Heat map of H3K4me1, H3K4me2, H3K27Ac and p300 within a −3 kb/+ 3 kb window centred on PU.1 in 18,405 macrophage-restricted enhancers in ES cells. Mφ, macrophage. **b**, Tag density of ATAC-seq and DNase–seq in 18,405 macrophage-restricted enhancers and random regions in ES cells shows chromatin openness of macrophage-restricted enhancers in ES cells. **c**, Target percentage of macrophage-restricted enhancers bound by ES cell transcription factors (ESRRB, NANOG, OCT4 and SOX2) within −1 kb/+ 1 kb. **d**, ENOS factor binding within a −1 kb/+ 1 kb window centred on each ES cell transcription factor in 6,775 premarked macrophage-restricted enhancers and 28,450 active ES cell-restricted enhancers. **e**, Binding of 12 ES cell transcription factors (OCT4, SOX2, NANOG, ESRRB, SMAD1, E2F1, TCFCP2L1, ZFX, STAT3, KLF4, C-MYC and N-MYC) in 6,809 active macrophage-restricted enhancers and 8,209 active ES cell-restricted enhancers defined by H3K27Ac (over 100 tags) in −1 kb/+ 1 kb windows. ChIP–seq data from published sources are listed in Supplementary Table 1.

To further explore any potential effect of ENOS binding in ES cells on the subsequent function of enhancers bound by PU.1 (also known as SPI1), a key macrophage transcription factor, we assessed the physical locations of ENOS binding in premarked enhancers. The locations of ENOS binding with respect to PU.1 binding sites varied in the 6,775 macrophage enhancers, corresponding to the locations of cognate binding sites (Fig. 2a).

Because enhancer RNA (eRNA) expression is a mark of enhancer activity[13], we investigated whether transcription units were present in ENOS-bound premarked enhancers. Because global run-on sequencing (GRO-seq) data on these enhancer regions were insufficiently robust to draw clear conclusions, we performed precision nuclear run-on sequencing for RNA polymerase II initiation sites (PRO-cap). We identified 2,336 significant RNA polymerase II initiation sites (cap sites) in −1 kb/+ 1 kb windows around 6,775 premarked enhancers in ES cells. While cap sites were found to be located close to PU.1 binding sites in macrophages, they were located at various distances from PU.1 sites in ES cells (Fig. 2b). The median distance from the PU.1 site to the macrophage-specific eRNA cap (−500 bp/+ 500 bp from the PU.1 site) was about 105 bp in mature macrophages[13], whereas the median distance from the ENOS binding site to the ES cell-specific eRNA cap was about 165 bp (Fig. 2c). The ES cell-induced transcription unit for each enhancer was therefore distinct from the eRNA transcription units subsequently nucleated by PU.1 and C/EBPα in mature macrophages, the binding sites for which are located within a core enhancer region.
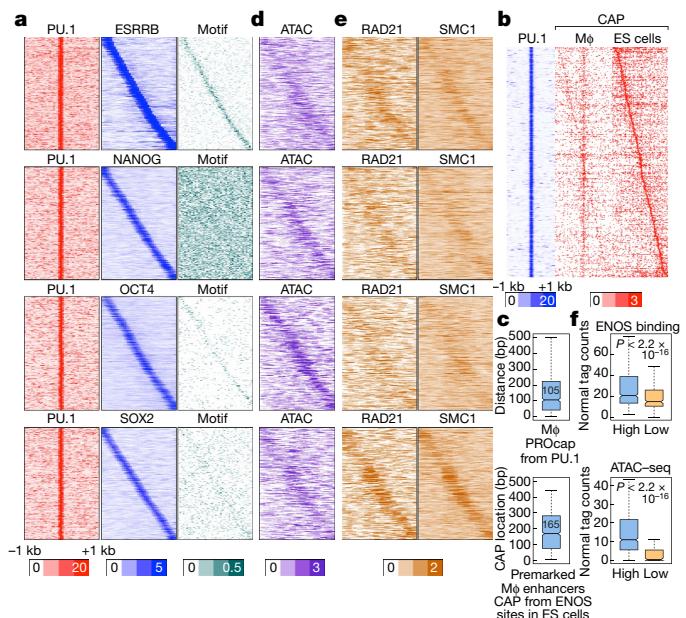
Fig. 2 | ES cell transcription factors bind to macrophage-restricted enhancers in ES cells. a, **d**, **e**, Heat map of ENOS binding (**a**), ATAC-seq (**d**), and RAD21 and SMC1 (**e**) in −1 kb/+ 1 kb windows in 6,775 premarked macrophage enhancers plotted relative to PU.1 binding sites in enhancers. **a**, Last panel shows corresponding binding sites of ENOS (ENOS motif) to the actual ENOS binding. **b**, Heat map of 2,345 PRO-cap peaks identified in macrophages or ES cells in −1 kb/+ 1 kb window from 6,775 premarked enhancers centred on PU.1. **c**, Locations of PRO-cap peak sites from PU.1 in macrophages or from ENOS in ES cells calculated in 638 premarked macrophage enhancers, with PRO-cap peaks in a −500 bp/+ 500 bp window from PU.1 binding sites. **f**, ENOS binding and ATAC-seq signal in highest versus lowest (high and low, respectively) approximately 40% of macrophage enhancers selected on the basis of RAD21 binding intensity in ES cells. In the box plots, line shows median and box shows 25th and 75th percentiles. P values calculated using Welch's two-sided t-test.

Indeed, ATAC-seq revealed chromatin openness in the region where ENOS bound to the enhancers (Fig. 2d). We investigated the functional correlation between cohesin and ENOS-bound enhancers, and found that high RAD21 binding was associated with high ENOS binding and a more open configuration (Fig. 2f).

To address any concerns regarding serum culture conditions, we performed chromatin immunoprecipitation with sequencing (ChIP–seq) with H3K4me2 and H3K27Ac, and ATAC-seq under two different culture conditions: with inhibitors of MEK and GSK3 (2i) and with serum. H3K4me2 and H3K27Ac were not observed in macrophage enhancers in either 2i- or serum-cultured ES cells, and the ATAC-seq signal was equivalently detected in both conditions (Extended Data Fig. 4a, b), indicating that culturing ES cells in serum did not affect our observations. Indeed, most pluripotency-associated genes were transcribed at similar levels in serum- and 2i-cultured ES cells[14] and only 8% of serum-treated ES cells were heterogeneous[15].

To investigate the potential function of premarking, we examined the correlation between ENOS factors and the ultimate activity of the enhancers in mature macrophages. ENOS-bound macrophage enhancers with the highest or lowest approximately 20% levels of ENOS transcription units were selected to test their ultimate activation in macrophages. The highest group of enhancers exhibited higher activation, as determined by binding of PU.1, PRO-cap signal, eRNA transcription, levels of H3K4me2 and H3K27Ac, and binding of cohesin in macrophages (Fig. 3a–d). In addition, when the highest or lowest approximately 20% of active macrophage enhancers were extracted using GRO-seq signal in macrophages[13], ENOS binding and ATAC-seq signals in ES cells were higher in the most active macrophage enhancers than in the least active macrophage enhancers (Fig. 3e, f).
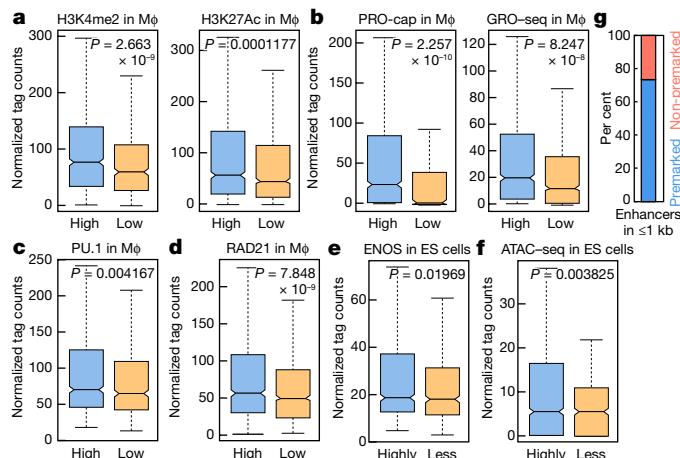
Fig. 3 | ES cell transcription factor binding is predictive of future enhancer activity. a–d, Comparison of the highest or lowest approximately 20% of enhancers selected on the basis of ENOS binding and PRO-cap signal in ES cells, and the levels of H3K4me2 and H3K27Ac (**a**), PRO-cap and GRO-seq (**b**), PU.1 (**c**) and RAD21 (**d**) in mature macrophages. **e, f**, Comparison of the highly active or less active about 20% of macrophage enhancers selected on the basis of GRO-seq signal of premarked macrophage enhancers in macrophages, and the levels of ENOS (**e**) and ATAC-seq (**f**) in ES cells. **g**, Percentage of functional macrophage enhancers in 6,775 enhancers that have ENOS binding in a −1 kb/+1 kb window (premarked enhancers) versus 11,630 enhancers that do not (non-premarked enhancers). In the box plots, line shows median and box shows 25th and 75th percentiles. P values calculated using Welch's two-sided t-test.

Indeed, putatively active macrophage enhancers, defined by eRNA transcription, were found much more frequently in the 6,775 ENOS-bound enhancers than in the 11,630 macrophage enhancers that do not bind ENOS in a −1 kb/+1 kb window (Fig. 3g). The correlation between premarking and enhancer robustness was confirmed in other tissues: spleen, lung, cortex and bone marrow (Extended Data Fig. 5a–d).

We next investigated whether the binding of these factors in ES cells had a direct role in the ultimate activation of cell-type-restricted enhancers in mature macrophages, choosing premarked macrophage-restricted enhancers linked to coding target genes that would not be predicted to influence macrophage development. We first selected a putative enhancer near the *Tlr1* locus, and then used CRISPR–Cas9 technology in ES cells to selectively delete the ESRRB binding site, followed by differentiation[16] (Extended Data Fig. 6b). Mature macrophages were selected on the basis of their ability to adhere to the non-adherent culture plates, as confirmed by expression of CD11B and F4/80 (Extended Data Fig. 6a). This permitted only a limited harvest of mature macrophages, thus precluding global genomic analyses. Notably, three independently derived individual clonal lines (#3, #10 and #14) were sequence-proven to harbour a 8-bp deletion of the ESRRB site in the *Tlr1* enhancer (Extended Data Fig. 6c, d). Consistent with the confirmed deletion, ESRRB binding in ES cells was decreased in these mutant clones compared to wild-type clones (Fig. 4a). The mutant clonal cells were differentiated into macrophages with an equivalent efficiency to that of wild-type cells; qualitative PCR with ChIP (ChIP–qPCR) for PU.1 and H3K4me2 was performed, and eRNA transcription was measured in the ES cell-derived macrophages (ESDMs). These analyses revealed that PU.1 binding was inhibited, that eRNA transcription was lost and that there was a consistent decrease in the level of H3K4me2 after deletion of the ESRRB site (Fig. 4b). To corroborate these findings, we constructed mutant clonal cells targeting the *Tnfaip3* enhancer with either a 16-bp deletion (#26) or a 21-bp deletion (#45) encompassing the ESRRB binding site (Extended Data Fig. 6e–g), and differentiated them into macrophages. The reduction in PU.1 recruitment, eRNA transcription and H3K4me2 levels
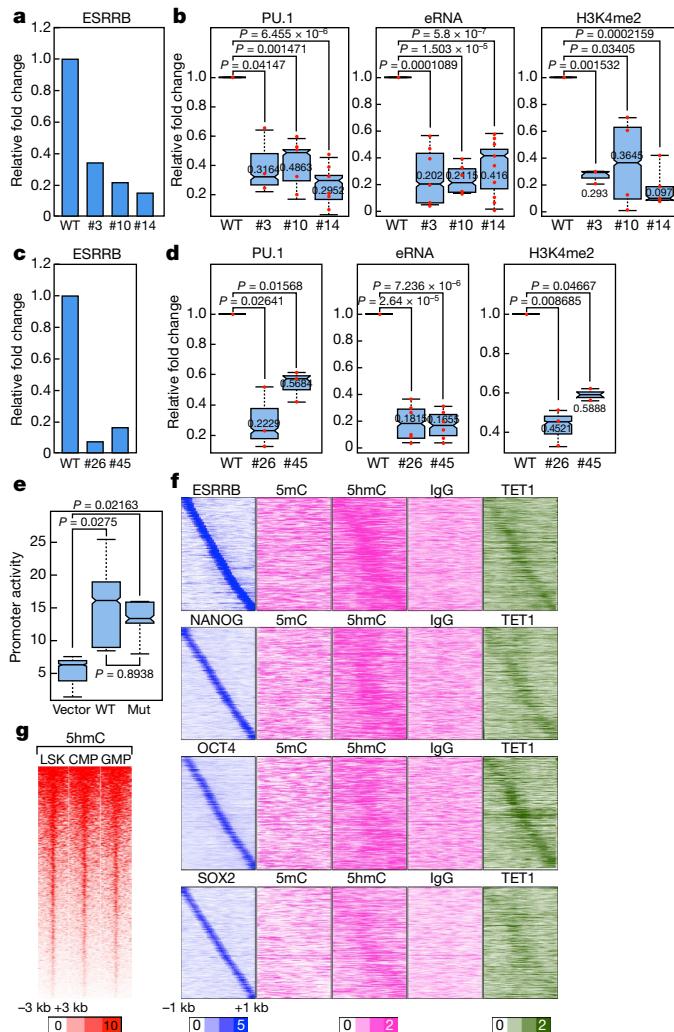
Fig. 4 | Premarking in ES cells is functionally required for robust future enhancer activation. a, c, Binding of ESRRB in wild-type (WT) and mutant clones of *Tlr1* (**a**; #3, #10, #14) and *Tnfaip3* (**c**; #26, #45) in ES cells. One example of representative data is shown ($n \geq 2$ biological repeats). **b, d**, PU.1 binding, eRNA transcription, and H3K4m2 level in wild-type and mutant clones of *Tlr1* (**b**) and *Tnfaip3* (**d**) in ESDMs. Each dot indicates a biological experiment ($n \geq 3$ biological repeats from two pooled different experiments, $n = 2$ biological repeats from two pooled different experiments for #45 for H3K4me2 ChIP–qPCR). **e**, Promoter activities in native full-length *Tlr1* enhancer response to wild-type versus ESRRB-deleted mutant in Raw264.7 cells ($n = 5$ biological repeats). **f**, Mapping of DNA methylation modification (5mC and 5hmC) and binding of TET1 in 6,775 premarked- macrophage enhancers in a −1 kb/+1 kb window centred on PU.1. **g**, 5-hmC in LSK (LT-HSC, ST-HSC, MPP: long term- and short term-hematopoietic stem cells, multipotent progenitors), CMP (common myeloid progenitors) and GMP (granulocyte macrophage progenitors) in 6,775 premarked enhancers in −3 kb/+3 kb window centred on PU.1. In the box plots, line shows median, and box shows 25th and 75th percentiles. P values calculated using Welch's two-sided t-test. Data from published sources are listed in Supplementary Table 1. Primer sets are listed in Supplementary Table 2.

in mutant ESDMs correlated with inhibition of ESRRB binding in ES cells (Fig. 4c, d). As a control, we tested *Tlr1* eRNA transcription in *Tnfaip3* mutant cells, observing no change (Extended Data Fig. 6h). We investigated two other loci, *Prdx5* and *Nod2*, in the same way (Extended Data Fig. 7a, b, e, f) and again found that enhancer activities defined by eRNA transcription were inhibited in mutant ESDMs with loss of ESRRB binding in ES cells, but *Prdx5* mutant enhancers (B11, F1, G11) exhibited the same level of *Nod2* eRNA transcription as in wild-type cells (Extended Data Fig. 7c, d, g). Because modifying the genomic

locus could potentially result in different genetic events, we performed RNA-seq in *Tlr1* enhancer clonal cells. Normal ES cells and clonal ES cells (wild-type, #10 and #14) exhibited the same pattern of transcription in the *Tlr1* locus (Extended Data Fig. 6i).

Enhancers harbouring the ESRRB core site deletion were evaluated by luciferase reporter assay to test whether genomic sequence disruption had impaired enhancer function compared to wild-type enhancers in macrophages. The mutant enhancers were equally competent to increase reporter expression as the wild-type enhancer in immortalized Raw 264.7 macrophages (Fig. 4e), confirming that the inhibitory effects in *Tlr1* mutant ESDMs resulted from inhibition of ESRRB binding in ES cells.

We further investigated this functional linkage in neural precursor cell (NPC)-restricted enhancers in the *Nek1* locus and *Ankrd1* locus (referred to as N4 and N8, respectively). Sox2 and Esrrb bound to N4 and Esrrb bound to N8, and we targeted the Sox2 or Esrrb binding site, in N4 or N8, respectively, for deletion (Extended Data Fig. 8a, d). Clonal ES cells were differentiated to NPCs in N2B27 medium for 6 days, with more than 70% of cells exhibiting the green signal caused by knock-in of the Sox1–GFP reporter in ES cells[17]. NPC differentiation was further confirmed by qRT–PCR of ES cell-specific genes (*Oct4*, *Esrrb* and *Nanog*) and NPC-specific genes (*Sox1*, *Fgf5*, *Nestin* and *Pax6*) (Extended Data Fig. 8g). Transcription of N4 or N8 eRNA was inhibited in N4- or N8-mutant ES cell-derived NPCs, respectively, correlating with inhibition of ESRRB binding in ES cells (Extended Data Fig. 8b, e). SOX2 binding could not be assessed because the anti-SOX2 IgG was ineffective. The specificity of N4- or N8-mutant clones on inhibition of eRNA transcription was confirmed by testing eRNA transcription in the N8 locus in N4 mutant cells, or the N4 locus in N8 mutant cells (Extended Data Fig. 8c, f). Collectively, these analyses corroborate the functional importance of ES cell factor premarking in both macrophage-restricted and NPC-restricted enhancers.

Poised enhancers, another class of well-studied cell-type-specific enhancer that is marked in ES cells, are often found near lineage-determining factors[18,19]. The 214 macrophage enhancers exhibit poised chromatin signatures (Extended Data Fig. 1a) with binding of about 2–4 ENOS factors in the enhancer cores in ES cells (Extended Data Fig. 9a, b). For example, we noted that *Spi1* (which encodes PU.1) has four clustered enhancers with poised chromatin signatures in ES cells. In contrast to the two enhancers (Enh2 and Enh3) that are active in mature macrophages, the other two (Enh1 and Enh4), which are bound by ESRRB and OCT4, are not (Extended Data Fig. 9c, d). These two premarked enhancers in ES cells might participate in the early activation of *Spi1*, which is known to exhibit a positive feedback loop on the two PU.1-bound enhancers[20].

A particularly intriguing question is how the marking of future cell-type-restricted enhancers might be 'remembered' for their ultimate activation later in development. Published results and analysis of the mRNA levels of the ENOS factors suggest that the temporal pattern of disappearance of these factors virtually coincides with the appearance of the first lineage-determining factors, TAL1, GATA2 and RUNX1, followed by the appearance of PU.1 and C/EBPα, which initiate macrophage enhancer activation[21] (Extended Data Fig. 10a). Another potential explanation is that binding of one of the ES cell transcription factors to the future cell-type-restricted enhancers might be accompanied by a specific DNA demethylation event that serves to ensure that an enhancer remains accessible to transcription factors, the binding of which may be impaired by DNA methylation. We therefore analysed available data regarding DNA methylation in ES cells, and found the 5-hydroxymethylcytosine (5-hmC) mark and the enzyme responsible for that mark, TET1, in the 6,775 premarked macrophage enhancers (Fig. 4f), reflecting the presence of TET1 in a complex with ESRRB and OCT4[22]. Notably, knockdown of *Esrrb* reduced 5-hmC levels in ESRRB-bound macrophage enhancers in ES cells (Extended Data Fig. 10b). To investigate whether 5-hmC is maintained during differentiation, such that it could serve as a marker for molecular memory, we examined 5-hmC during haematopoiesis using published data[23],

and found that 5-hmC was maintained during haematopoiesis (Fig. 4g, Extended Data Fig. 10c). The enhancer histone marks, H3K4me1 and H3K27Ac, were studied during haematopoiesis[24], but not found in ES cells, or even in mesoderm, being gained gradually early in haematopoiesis (Extended Data Fig. 10c, d).

Investigation of the signature of cell-type-restricted enhancers in ES cells has provided an insight into the process of genomic enhancer recognition in ES cells underlying cell-type-specific transcriptional programs. Furthermore, the marking of the cell-type-restricted enhancers in ES cells license the ultimate robust activation of the cell-type-restricted enhancer in mature differentiated cells (Extended Data Fig. 10e).

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0048-8.

1. Heintzman, N. D. et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
2. Miguel-Escalada, I., Pasquali, L. & Ferrer, J. Transcriptional enhancers: functional insights and role in human disease. *Curr. Opin. Genet. Dev.* **33**, 71–76 (2015).
3. Ong, C.-T. & Corces, V. G. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* **12**, 283–293 (2011).
4. Kaikkonen, M. U. et al. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* **51**, 310–325 (2013).
5. Huang, H.-T. et al. A network of epigenetic regulators guides developmental haematopoiesis in vivo. *Nat. Cell Biol* **15**, 1516–1525 (2013).
6. Epelman, S., Lavine, K. J. & Randolph, G. J. Origin and functions of tissue macrophages. *Immunity* **41**, 21–35 (2014).
7. Chen, X. et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
8. Boyer, L. A. et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
9. Schmidt, D. et al. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578–588 (2010).
10. Phillips-Cremins, J. E. et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
11. Kagey, M. H. et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
12. van de Werken, H. J. G. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972 (2012).
13. Li, W. et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).
14. Ying, Q.-L. et al. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
15. Kumar, R. M. et al. Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature* **516**, 56–61 (2014).
16. Zhuang, L. et al. Pure populations of murine macrophages from cultured embryonic stem cells. Application to studies of chemotaxis and apoptotic cell clearance. *J. Immunol. Methods* **385**, 1–14 (2012).
17. Ying, Q.-L., Stavridis, M., Griffiths, D., Li, M. & Smith, A. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.* **21**, 183–186 (2003).
18. Wang, A. et al. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16**, 386–399 (2015).
19. Rada-Iglesias, A. et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
20. Will, B. et al. Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. *Nat. Med.* **21**, 1172–1181 (2015).
21. Lichtinger, M. et al. RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J.* **31**, 4318–4333 (2012).
22. Gagliardi, A. et al. A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. *EMBO J* **32**, 2231–2247 (2013).
23. Han, D. et al. A highly sensitive and robust method for genome-wide 5hmC profiling of rare cell populations. *Mol. Cell* **63**, 711–719 (2016).
24. Lara-Astiaso, D. et al. Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943–949 (2014).

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0048-8.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0048-8.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to H.S.K. or M.G.R.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Cell culture. 46c mouse ES cells were a gift from A. Smith. Cells were grown in feeder-free conditions as previously described[17]. ES cells were maintained in serum culture medium with DMEM-KO (Invitrogen 10829-018) supplemented with 15% ES cell qualified-fetal bovine serum (Omega, FB-05), 2 mM nonessential amino acids (Invitrogen 11140-050), glutamax (Invitrogen 35050061), penicillin/streptomycin (Invitrogen 15140122), 2-mercaptoethanol (Sigma, M7522) and 1,000 U/ml LIF (ESGRO, ESG1106). ES cells in 2i medium were grown in N2/B27 medium with 50% neurobasal (Gibco 21103-049) and 50% DMEM/F12 (Invitrogen 21331-020), 2 mM nonessential amino acids, glutamax, penicillin/streptomycin, 2-mercaptoethanol, N2 supplement (Invitrogen, 175020-01), B27 (Invitrogen, 17504-001), 1,000 U/ml LIF, and 2i (ESGRO, ESG1121). Cells were regularly checked for mycoplasma contamination (Lonza, LT07-318). Peritoneal macrophages were obtained from 6–8-week-old female C57BL/6 J mice (Jackson laboratory). Mice were injected with thiogycollate for 3–4 days, and macrophages were obtained and cultured in RPMI supplemented with 10% FBS and penicillin/streptomycin overnight before collection. All animal care and experimental procedures were in accordance with the University of California, San Diego research guidelines for the care and use of laboratory animals.

Antibodies. Antibodies used for ChIP and ChIP–seq included: RAD21 (Abcam, ab992), H3K4me2 (Upstate, 07-030), PU.1 (Santa Cruz, sc-352) and H3K27Ac (Active Motif, 39133); for immunostaining, F4/80-FITC (eBioscience, 11-4801-81) and CD11b-PE-Cyanine5 (eBioscience, 15-0112-81); for MeDIIP, 5-hmC (Active Motif, #39769).

Enhancer reporter assay. For construction of *Tlr1* enhancer reporter plasmids, 600 bp of *Tlr1* locus spanning the ESRRB binding site was PCR amplified and cloned into pGL4.23 at the KpnI/XhoI sites downstream of the luciferase reporter gene controlled by a minimal promoter. Enhancer reporters were transfected into RAW264.7 macrophages using Lipofectamine 3000 (Invitrogen), using 200 ng of enhancer reporter and 5 ng of Renilla. Luciferase activity was measured 48 h after transfection using a Veritas microplate luminometer (Turner Biosystems) and normalized to Renilla activity to correct for differences in transfection efficiency.

ES cell differentiation to macrophages. ES cell differentiation to macrophages was performed as previously described[16]. In brief, ES cells were trypsinized and transferred to bacteriological plates in macrophage differentiation medium, which has 15% L929 conditioned medium and 1 ng/ml IL3 but lacks LIF, for about 6–8 days to make embryoid bodies (EBs). EBs were transferred onto gelatin-coated tissue culture dishes for about 3–4 days. After that, supernatants of adherent EBs containing floating macrophage progenitors were collected and plated onto bacteriological dishes for 7 days to obtain adherent macrophages in macrophage differentiation medium, which has 15% L929 conditioned medium. Medium was readded to the adherent EB plates and macrophage progenitors were obtained every 2 days.

ES cell differentiation to NPCs. ES cell differentiation to NPCs was performed as previously described[17]. In brief, ES cells were trypsinized, plated onto gelatin-coated plates and cultured in N2B27 medium for 6 days. On day 6, > 70% green-signal expressing cells were observed, which indicates differentiation into NPCs, and cells were collected for further experimental analysis.

CRSIPR–Cas9 assay. Online software (http://crispr.mit.edu) was used to design optimal candidate sgRNAs to target proximal regions of ES cell transcription factor motifs, and these sgRNAs were cloned to vector PX459 (Addgene #48139), which co-expresses sgRNA and Cas9. This plasmid was transfected into ES cells with Lipofectamine 2000 and 1.0 μg/ml puromycin was added 2 days after transfection. Cells were cultured for another 3 days then diluted to pick up single-cell-derived clonal lines. DNA was extracted from these cells and PCR was performed to amplify fragments containing sgRNA-targeted ES cell transcription factor (ESRRB or SOX2) motifs. Sequencing was applied to identify clones in which the ESRRB or SOX2 motif was mutated or deleted. sgRNA sequences and primer sequences are listed in Supplementary Table 2.

RNA preparation and qRT–PCR. RNA was isolated using Trizol (Invitrogen) or RNeasy kit (Qiagen). RNA was reverse transcribed using SuperScript III Reverse Transcriptase kit (Invitrogen). Quantitative PCR was performed using SYBR Green Master Mix (Biorad). The experiments were repeated at least three times as biological replicates and *P* values were obtained using Welch's two-sided *t*-test. Primers are listed in Supplementary Table 2.

RNA-seq. RNA-seq libraries were made from 1 μg DNase-treated total RNA using TruSeq stranded mRNA Library Prep Kit (Illumina, RS-122-2101, RS-122-2102) according to the manufacturer's instructions. RNA-seq libraries were quantified with Qubit, clustered and sequenced using HiSeq 3000/4000 SR cluster kit (Illumina GD-410-1001) and HiSeq 3000/4000 SBS kit (Illumina FC-410-1001). RNA-seq reads were counted using HOMER software considering only exonic regions for RefSeq genes. The proximal macrophage-expressed genes of 6,775 macrophage-restricted enhancers were examined and further analysis was performed to obtain differentially expressed genes using the following criteria: FC ≥ 4, over 20 tags in macrophage, and visualized by heat map with $\log_2$ transform.

Gene ontology analysis and genetic association analysis were performed using Metascape (http://metascape.org).

PRO-cap. Pro-cap and library preparation for sequencing have previously been described[25]. Nuclei were prepared from ~40 million cells for run-on assay. Run-on reactions were stopped and RNA was extracted with Trizol LS reagent (Invitrogen). Following DNase treatment, the RNA was fragmented. Biotin-incorporated fragmented RNA was immunoprecipitated with anti-strepavidin beads (Invitrogen). Then the RNA was treated with polynucleotide kinase (Enzymatics) and precipitated. The RNA was dephosphorylated with calf intestinal phosphatase (NEB) and 5′-de-capped with tobacco acid pyrophosphatase (Epicentre). The reaction was stopped and RNA was extracted with Trizol LS, and libraries were prepared by ligating Illumina TruSeq-compatible adapters to the RNA 3′ and 5′ ends with truncated mutant RNA ligase 2 (K227Q) and RNA ligase 1 (NEB), respectively, followed by reverse transcription, cDNA isolation and PCR amplification. Final libraries were size selected on TBE gels to 60–110-bp insert size. Pro-cap results were trimmed to remove A-stretches originating from the library preparation. Each sequence tag returned by the Illumina Pipeline was aligned to the mm9 assembly using Bowtie2 allowing up to three mismatches. Only tags that mapped uniquely to the genome were considered for further analysis. Each sequencing experiment was normalized to a total of 10^7 uniquely mapped tags by adjusting the number of tags at each position in the genome to the correct fractional amount given the total tags mapped.

ChIP–seq. Cells were fixed with 2 mM disuccinimidyl glutarate (DSG) (proteochem) for 45 min and 1% formaldehyde for 10 min, and quenched using glycine for 5 min. Nucleus lysates were prepared using sonication buffer with 1% SDS, and immunoprecipitation was performed with several different antibodies. After overnight incubation with antibodies, beads were added for another 3 h, and washing was performed. Reverse-crosslinking was done overnight at 65 °C and DNA was purified using QIAquick Spin column (Qiagen). For ChIP–seq, extracted DNA was ligated to adaptors and deep sequencing was performed with Illumina's HiSeq 2000, 2500, or 4000 system according to the manufacturer's instructions. The first 48 bp of each sequence tag returned by the Illumina Pipeline was aligned to the mm9 assembly using BFAST or Bowtie2. Only uniquely mapped tags were selected for further analysis. The data were visualized by preparing custom tracks on the UCSC genome browser using HOMER. Genomic binding peaks for transcription factors were identified using the findPeaks.pl command from HOMER with eightfold enrichment over the input sample, fourfold enrichment over local background, a minimal tag number of 16, and normalization to 10^7 mapped reads per experiment. For histone marks, initial seed regions of 500 bp were considered to calculate enriched reads.

Regions of maximal density exceeding a given threshold with FDR < 0.001 were called as peaks. Peaks within ± 1,000 bp of from the RefSeq gene transcription start site were considered to be promoters, and to focus the analysis on enhancers, peaks within 3 kb of a gene promoter were filtered out.

4C–seq. Chromosome confirmation capture was performed as previously described[12]. In brief, 10 million cells were cross-linked with 1% formaldehyde for 10 min and quenched with glycine for 5 min on ice. Soluble chromatin was incubated with 400 U HindIII (NEB) overnight at 37 °C, and then intramolecular ligations were performed with 1,000 U T4 DNA ligase (NEB) for 4 h at 16 °C under dilution. Chromatin was decrosslinked at 65 °C and then treated with RNase to remove RNA before purification using several phenol and phenol–chloroform extractions and ethanol precipitation. The second restriction digestion was also performed overnight at 37 °C, using 50 U DpnII (NEB). DNA was ligated overnight and purified as before and ultimately using Qiagen columns and subjected to inverse PCR (Expand Long-Range PCR system; Roche Diagnostics) using a first primer designed on the viewpoint near a HindIII site and a second outer primer designed beside the DpnII site. Both primers contained Illumina sequencing adapters and barcodes for multiplexing. PCR samples were purified using a Roche kit and quantified using a Qubit.

We analysed data using a previously described bioinformatics pipeline[12].

MeDIP–seq. Genomic DNA for hMeDIP was isolated from cells using Qiagen's DNeasy Blood and Tissue Kit. Isolated genomic DNA was then fragmented to 100–300 bp through sonication using Diagenode's Bioruptor platform and the size distribution was confirmed through gel electrophoresis. Barcoded adaptors for Illumina sequencing were added to 1 μg of fragmented genomic DNA per experiment, using the NEBNext Ultra II DNA Library Prep Kit for Illumina, following the manufacturer's instructions. This protocol was stopped after adaptor ligation and cleanup (and before any amplification steps), and the adaptor ligated fragmented DNA was then used for hydroxy-methylated DNA pulldown. Denaturing, immunoprecipitation, washing, and purification of hydroxy-methylated DNA were performed as previously described[26], with the following modifications. One microlitre of 5-hmC antibody-containing serum was used per immunoprecipitation reaction. Washing was done five times, with each wash for 15 min at 4 °C. Next, DNA was eluted from beads using 200 μl digestion buffer, incubated overnight

with Proteinase K at 50 °C, and purified using Qiagen's QIAquick PCR Purification Kit. Purified hydroxy-methylated genomic DNA was then further processed with NEBNext Ultra II DNA Library Prep Kit for Illumina at the step of PCR enrichment of adaptor-ligated DNA, per the manufacturer's instructions, continuing with the entire protocol to prepare libraries for Illumina sequencing. Data were mapped to mm9 using Bowtie2 with standard settings, and peaks were found by using MACS with default parameters.

ATAC-seq. ATAC-seq was performed as previously described[27]. Nuclei were prepared from 50,000 cells and the transposition reaction was performed for 30 min at 37 °C. DNA fragments were amplified by PCR and purified, and deep sequencing was performed with Illumina's HiSeq 2000, 2500, or 4000 system according to the manufacturer's instructions. ATAC-seq data were mapped to mm9 using Bowtie2 with standard settings. Tag directories with reads mapped to the mitochondrial chromosome filtered out were created. ATAC-seq peaks were identified using findPeaks.pl in HOMER with the settings: -style histone -size 75 -minDist 75 -minTagThreshold 6 -L 8 -F 8. BED files were created from Tag directories using the HOMER package.

Deep sequencing. For all ChIP–seq, RNA-seq, 4C–seq, ATAC-seq, MeDIP–seq and PRO-cap experiments, the DNA libraries were sequenced for 50 cycles on Illumina's HiSeq 200, 2500 or 4000 system according to the manufacturer's instructions. Sequencing experiments were visualized by preparing custom tracks for the UCSC Genome browser.

Bioinformatic characterization of enhancers. The criteria for identifying PU.1-H3K4me2 co-bound enhancer regions was that the distance from the centre of a PU.1 peak to the H3K4me2 peak-occupied region was ≤1 kb. ES cell factor-bound macrophage enhancers were defined by calculating the distance ≤1 kb between ES cell factor peak-spanning regions and PU.1 bound macrophage enhancers. The active ES cell enhancers or active macrophage enhancers were defined using H3K27Ac (over 100 tags) to examine binding of 12 different ES cell factors. The functional macrophage enhancers were defined using GRO-seq (over 20 tags) in macrophages, and these enhancers were used to count the number of functional enhancers in ENOS-bound enhancer or non-bound enhancer using a −1 kb/+1 kb window. Highly active or less active macrophage enhancers were created after excluding non-active macrophage enhancers using GRO-seq (less than 5 tags) in macrophages. The comparison of tag intensity of ChIP–seq, ATAC-seq, PRO-cap and GRO-seq or distances between different categories are presented as boxplots by using normal scales. P values were calculated using Welch's two-sided t-test. To profile the distribution of ES cell factors surrounding PU.1-H3K4me2 co-bound enhancer regions, ChIP–seq signals surrounding PU.1 peak centres were separated into 40 bins, and then were sorted by the tag numbers based on the distance to the PU.1 peak centre.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

GEO data. The GEO datasets used in this study are listed in Supplementary Table 1.

Data availability. All deep sequencing data have been deposited in GEO under accession number GSE81681.

25. Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).
26. Thu, K. L. et al. Methylated DNA immunoprecipitation. *J. Vis. Exp.* **23**, e935 (2009).
27. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
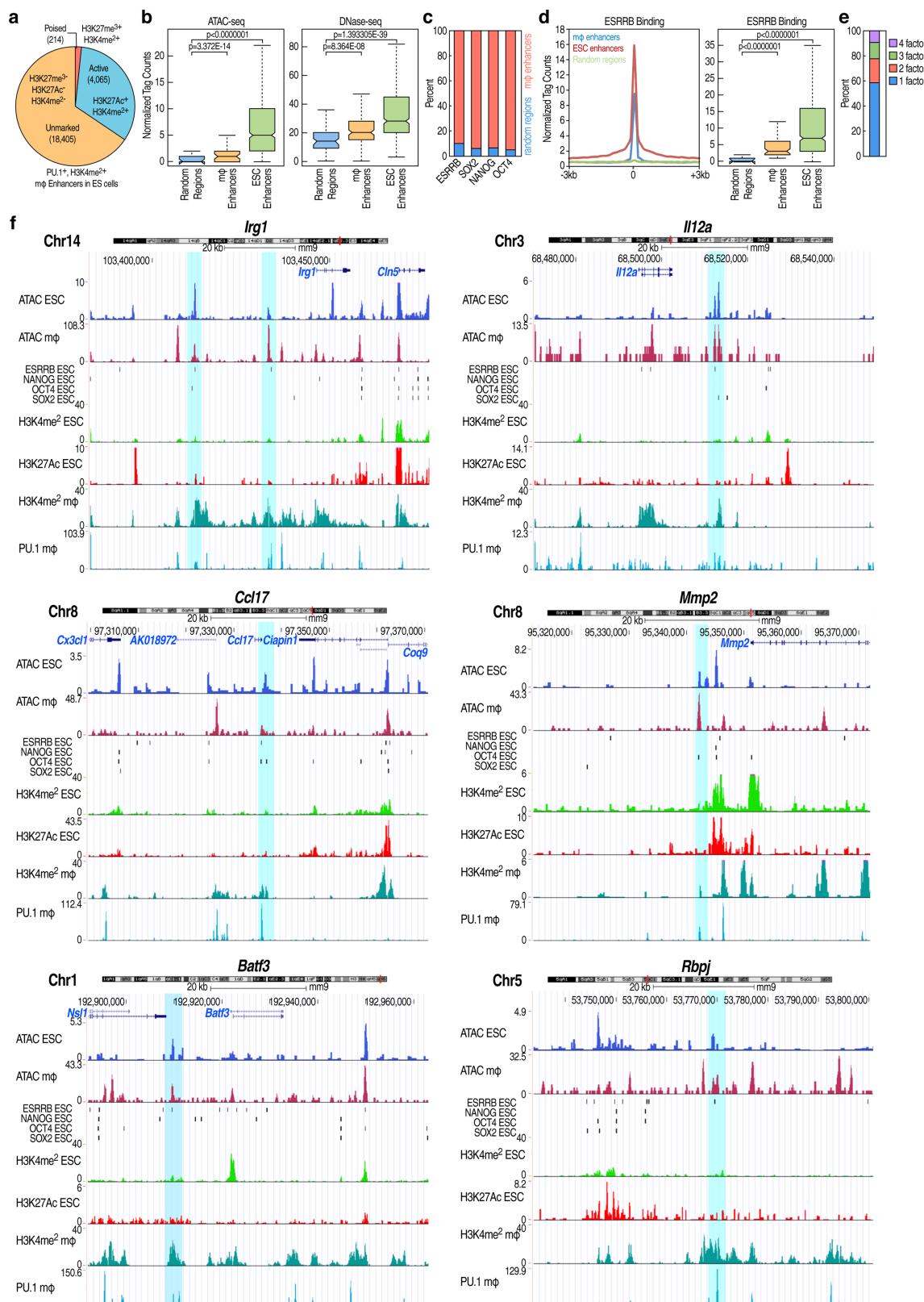
Extended Data Fig. 1 | Features of macrophage-restricted enhancers in ES cells. a, PU.1[+]H3K4me2[+] macrophage enhancer profiles in mouse ES cells. **b**, Normalized tag counts of ATAC-seq and DNase-seq in 18,405 random regions versus 18,405 macrophage enhancers versus 28,450 ES cell enhancers. **c**, Percentage of macrophage enhancers or random regions bound by ENOS in a −1 kb/+1 kb window. **d**, ESRRB binding in random regions versus macrophage enhancers versus ES cell enhancers in ES cells. **e**, ENOS binding profiles in 6,775 macrophage enhancers.

f, UCSC genome browser screen shots show ENOS binding with chromatin openness but no enhancer marks, such as H3K4me2 and H3K27Ac, in macrophage-restricted enhancers in ES cells. The *Irg1* gene locus has two different macrophage-restricted enhancers bound by ES cell transcription factors. The blue bar indicates macrophage-restricted enhancers. In the box plots, line shows median, and box shows 25th and 75th percentiles. *P* values calculated using Welch's two-sided *t*-test.
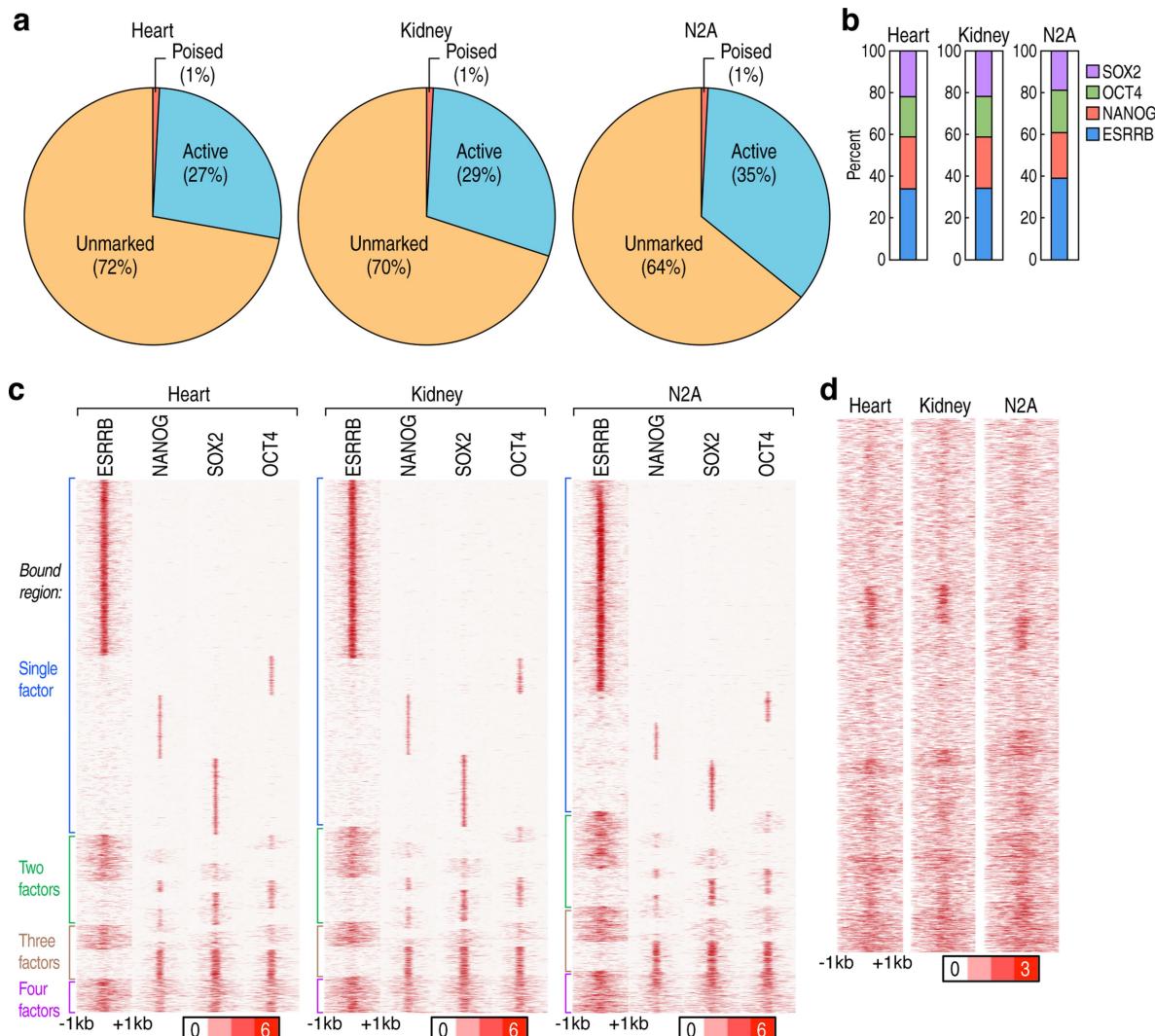
Extended Data Fig. 2 | ES cell transcription factors bind to heart-, kidney- and N2A cell-restricted enhancers in ES cells. a, Profiles of cell-type-restricted enhancers in heart, kidney, and N2A neuronal cells, with percentages of enhancers that are unmarked, poised and active in ES cells. **b**, Target percentage of ENOS-bound heart-, kidney- and N2A cell-restricted enhancers in ES cells. **c**, Heat map of ESRRB, NANOG, OCT4 and SOX2 binding within a −1 kb/+1 kb window centred on each transcription factor in 17,561 heart-, 14,242 kidney- and 12,027 N2A cell-restricted enhancers in ES cells. **d**, Heat map of ATAC-seq within a −1 kb/+1 kb window centred on each transcription factor in heart-, kidney- and N2A cell-restricted enhancers in ES cells. Data from published sources are listed in Supplementary Table 1.
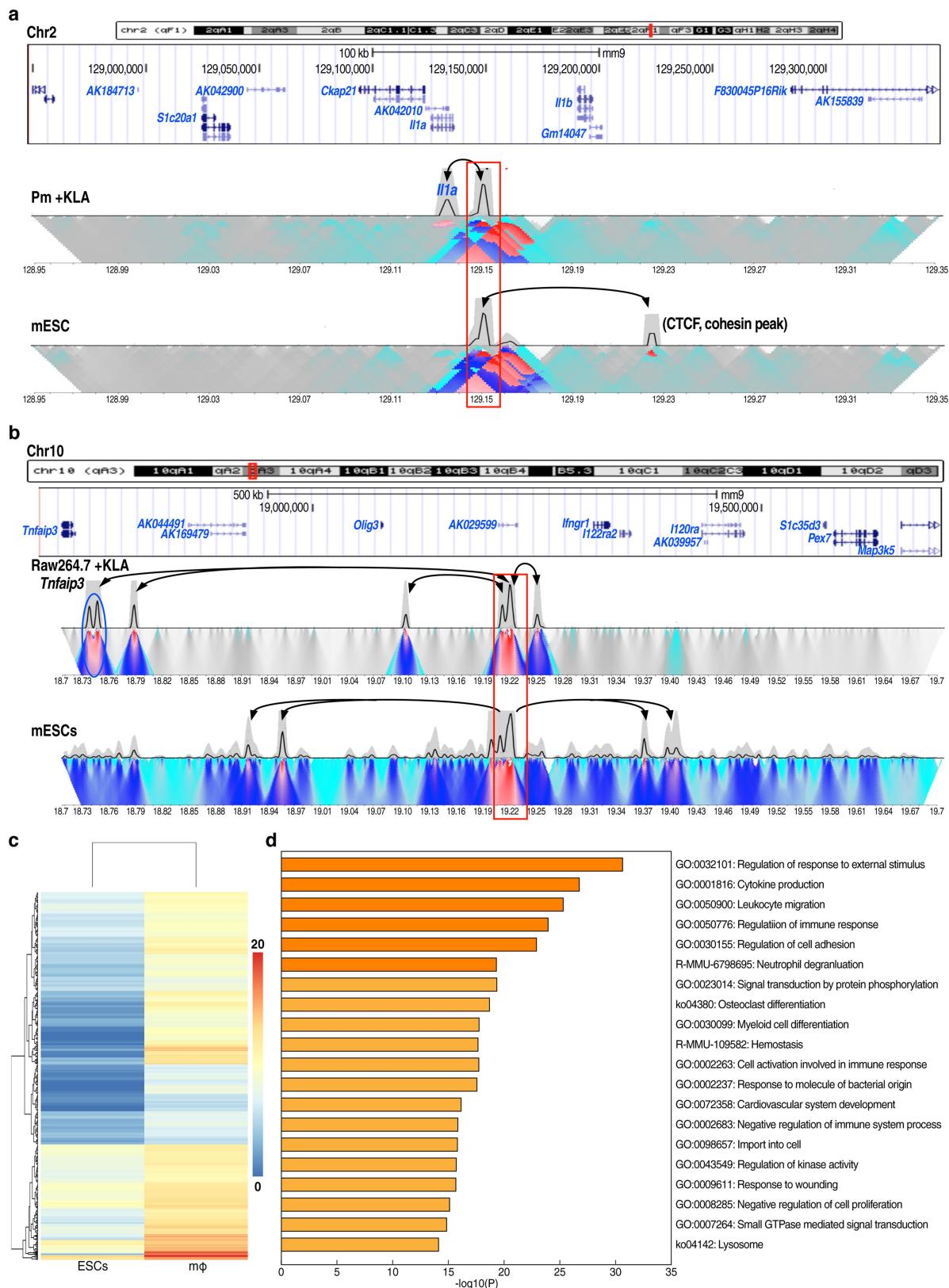
Extended Data Fig. 3 | Premarked macrophage-restricted enhancers do not interact with cognate target coding genes of macrophages in ES cells. a, b, 4C–seq in ES cells and macrophages. Primer sets for 4C are listed in Supplementary Table 2. **a**, 4C–seq analysis of the *Il1a* gene locus from the enhancer viewpoint in KLA-treated peritoneal macrophages or ES cells. Black arrows represent interaction frequency based on the trunc mean in a 400-kb window. The statistic of trunc mean calculated standard mean but with a truncated high and low value[12]. Red box shows position of macrophage putative enhancer 5′ of *Il1a*. **b**, 4C–seq analysis of the *Tnfaip3* gene locus from the enhancer viewpoint in KLA-treated

Raw 264.7 macrophages or in ES cells. Black arrows indicate interactions based on linear mean in 1-Mb window. Red box indicates position of putative *Tnfaip3* enhancer and blue circle indicates *Tnfaip3* promoter. **c**, The 634 differentially expressed genes (fold-change ≥ 4, over 20 tags in macrophage) among the proximal macrophage-expressed genes of 6,775 macrophage-restricted enhancers were visualized by heat map in ES cells and macrophages. Colour indicates normalized tag counts with $\log_2$ transform. **d**, Gene ontology (GO) term analysis of 634 differentially expressed genes.
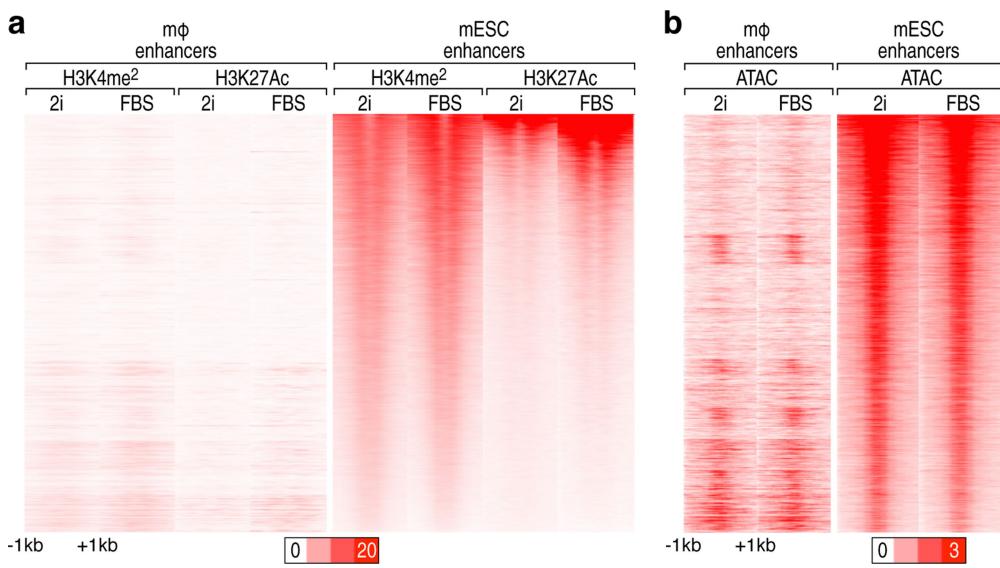
Extended Data Fig. 4 | Features of 6,775 macrophage enhancers in ES cells are not affected by ES cell culture medium (2i or serum). **a**, **b**, Heat map of H3K4me2 and H3K27Ac (**a**) and ATAC-seq (**b**) in 6,775 premarked macrophage enhancers and 28,450 ES cell enhancers in 2i- or serum-cultured ES cells. Centred on ENOS and visualized in a −1 kb/+1 kb window.
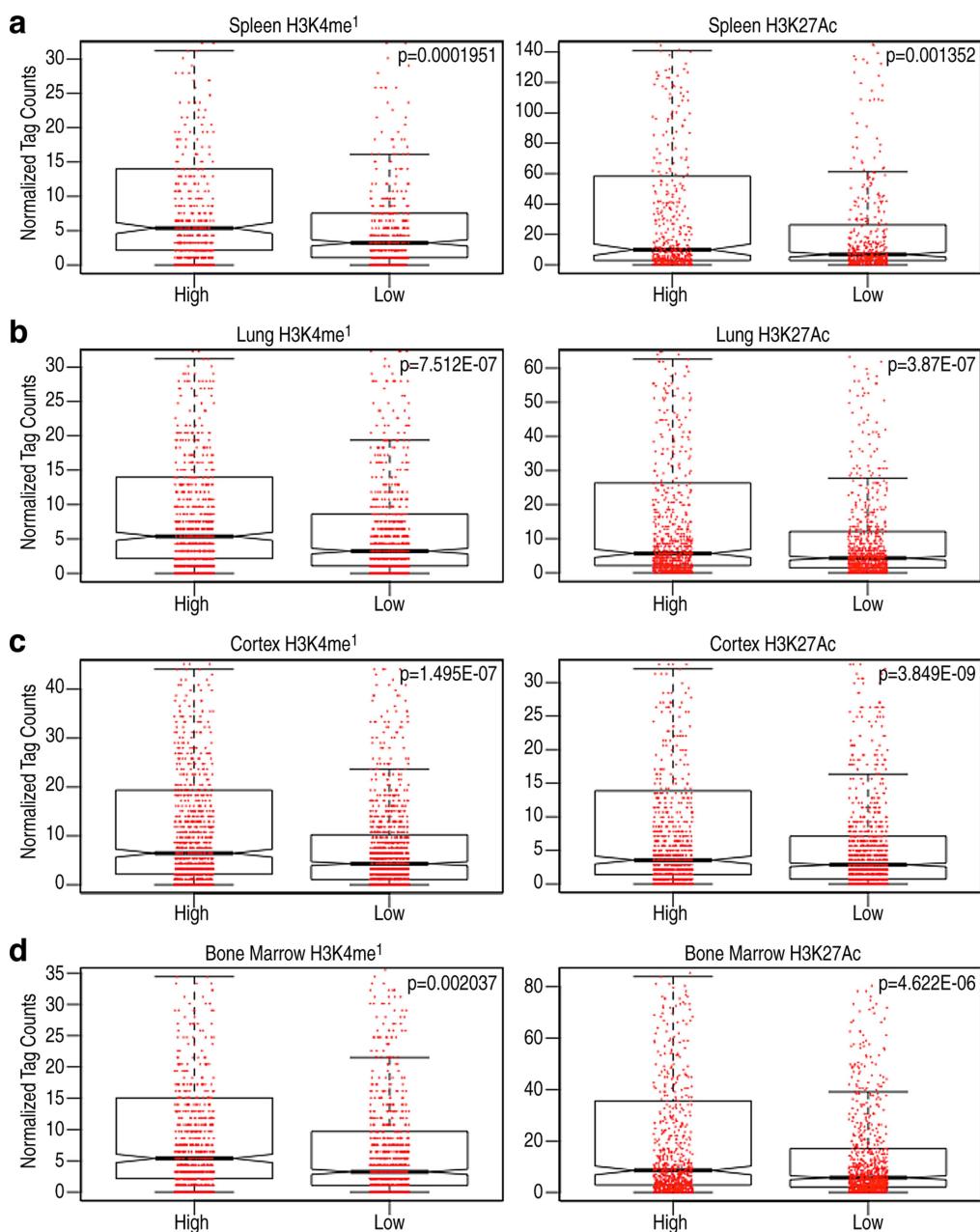
Extended Data Fig. 5 | ENOS binding correlates with future enhancer activity. a–d, Comparison of the approximately 10% 'most' marked (high) and 'least' marked (low) enhancers selected on the basis of ENOS binding, and the level of H3K4me1 and H3K27Ac on these enhancers in mature tissues for spleen (**a**), lung (**b**), cortex (**c**) and bone marrow (**d**). In the box plots, line shows median and box shows 25th and 75th percentiles. Red dots indicate each value of H3K4me1 or H3K27Ac. *P* values calculated using Welch's two-sided *t*-test. Data from published sources are listed in Supplementary Table 1.
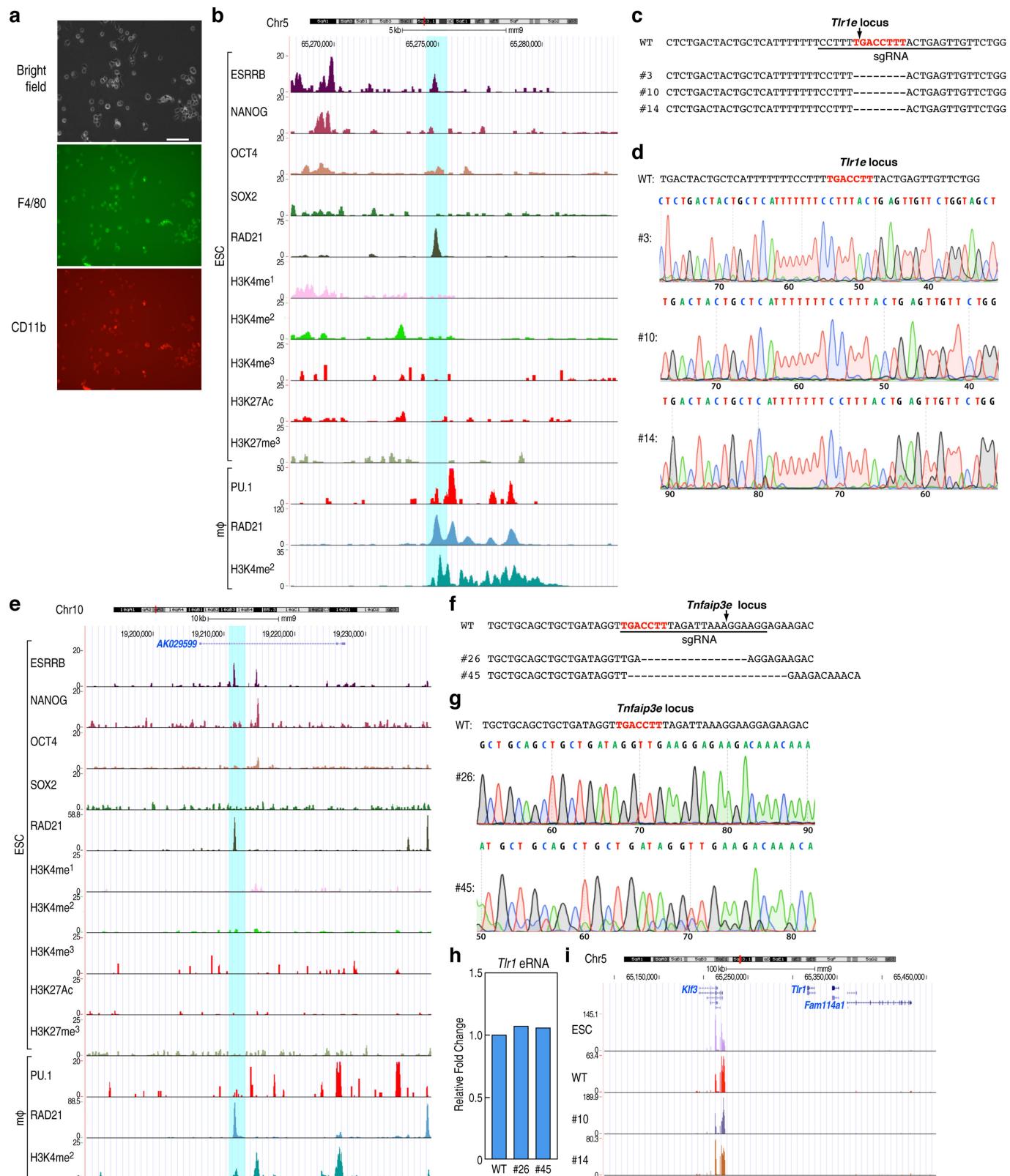
Extended Data Fig. 6 | Premarking in ES cells is required for robust future macrophage enhancer activation. a, Photomicrographs of macrophages differentiated from ES cells documented by F4/80 and CD11B staining. Scale bar, 50 μM. **b**, Screenshot of the *Tlr1* locus. The blue box corresponds to the CRISPR–Cas9 target region. **c, d**, Putative *Tlr1* enhancer mutant sequence diagram and DNA sequence documentation of several of the homozygous mutation clones, used for analysis. **e**, Screenshot of the *Tnfaip3* locus. The blue box corresponds to the

CRISPR–Cas9 target region. **f, g**, Putative *Tnfaip3* enhancer mutant sequence diagram and DNA sequence documentation of several of the homozygous mutation clones, used for analysis. **h**, *Tlr1* eRNA transcription was tested in wild-type and mutant clonal cells (#26, #45) of the *Tnfaip3* enhancer. One representative experiment is plotted (*n* = 2 biological repeats). **i**, UCSC genome browser shot of RNA-seq in *Tlr1* locus in normal ES cells and clonal mutant ES cells (wild-type, #10, #14). Primer sets are listed in Supplementary Table 2.
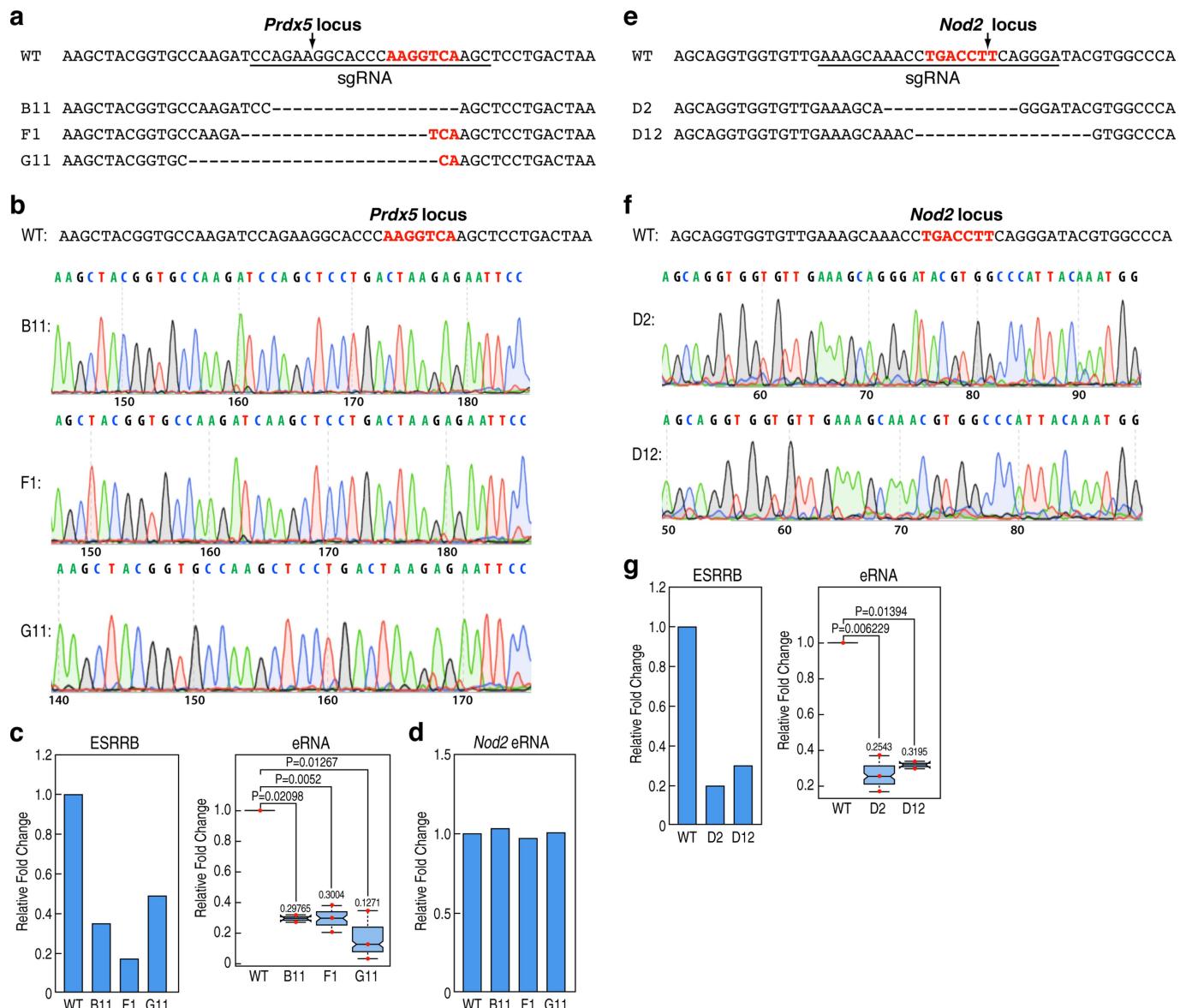
Extended Data Fig. 7 | Premarking in ES cells is required for robust future macrophage enhancer activation. a, b, Putative *Prdx5* enhancer mutant sequence diagram and DNA sequence documentation of several of the homozygous mutation clones, used for analysis. **c,** Three mutant clones (B11, F1, G11) of the putative *Prdx5* enhancer decreased level of *Prdx5* eRNA transcription in ESDMs (right). Each dot indicates one biological experiment (*n* = 3 biological repeats from two pooled different experiments; *n* = 2 biological repeats from two pooled different experiments for B11). ESRRB binding in ES cells is a representative experiment (left) (*n* = 2 biological repeats). **d,** *Nod2* eRNA transcription was tested in wild-type and mutant (B11, F1, G11) clonal cells of the *Prdx5* enhancer. Data from one representative experiment are plotted (*n* = 2

biological repeats). **e, f,** Putative *Nod2* enhancer mutant sequence diagram and DNA sequence documentation of several of the homozygous mutation clones, used for analysis. **g,** Two mutant clones (D2, D12) of the putative *Nod2* enhancer decreased level of *Nod2* eRNA transcription in ESDMs (right). Each dot indicates one biological experiment (*n* = 3 biological repeats from two pooled different experiments; *n* = 2 biological repeats from two pooled different experiments for D12). ESRRB binding in ES cells is a representative experiment (left) (*n* = 2 biological repeats). In the box plots, line shows median and box shows 25th and 75th percentiles. *P* values calculated using Welch's two-sided *t*-test. Primer sets are listed in Supplementary Table 2.

a

N4-WT   AGTGAGAGGAAAACCCCAGG**CACAATGG**GGGTTTAGTGGGTAACAAAAGGATGAATGAAATTCCTGTAGAA
         GATAAAGTGGGTGGAAGGCCAGAGGTCATTTAA

N4-C3   AGTGAGAGGAAAA--------------------------------------GGATGAATGAAATTCCTGTAGAA
         GATAAAGTGGGTGGAAGGCCAGAGGTCATTTAA

N4-C16   AGTGAGAGGA--TGAATGAAATTCCTGTAGAA
         GATAAAGTGGGTGGAAGGCCAGAGGTCATTTAA

N4-C3

N4-C16

b

ESRRB

N4 Enhancer RNA Transcription

c

N8 eRNA

d

N8-WT   TAGTCACACTGGGCGGGGC**AGGGTCA**CCGGAGCTTTGGGA**AAGGACA**GCTGTCCTCTTTTCTAGGTAAATGATGT

N8-C6   TAGTCACACTGGGCGGG---CTAGGTAAATGATGT
         TAGTCACACTGGGCGG--TTCTAGGTAAATGATGT

N8-C12   TAGTCACA--TCTAGGTAAATGATGT
         TAGTCTTG--GCAAGGACAGTAAATGATGT

N8-C6

N8-C12

e

ESRRB

N8 Enhancer RNA Transcription

f

N4 eRNA

g

ESC-expressing genes

NPC-expressing genes

Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Premarking in ES cells is required for robust future NPC enhancer activation. a, The putative *Nek1* enhancer, referred to as N4; mutant sequence diagram and DNA sequence documentation of several mutation clones (N4C3, N4C16), used for analysis. **b**, Two mutant clones (N4C3, N4C16) of the N4 enhancer decreased level of N4 eRNA transcription in NPCs (right). Each dot indicates one biological experiment. A representative experiment for ESRRB binding in ES cells is presented (left) ($n = 3$ biological repeats for N4C3, $n = 2$ biological repeats for N4C16). **c**, N8 eRNA transcription was tested in N4 mutant clones (N4C3, N4C16) to show specificity of N4 mutant cells on inhibition of N4 eRNA transcription. A representative experiment is presented ($n = 2$ biological repeats). **d**, The putative *Ankrd1* enhancer, referred to as N8; mutant sequence diagram and DNA sequence documentation of several of the mutation clones (N8C6, N8C12), used for analysis. **e**, Two mutant clones (N8C6, N8C12) of the N8 enhancer decreased level of N8 eRNA transcription in NPCs (right). Each dot indicates one biological experiment. ESRRB binding in ES cells is a representative experiment showing inhibition of ESRRB binding in ESRRB-deleted N8 locus (left) ($n = 3$ biological repeats). **f**, N4 eRNA transcription was examined to prove the specificity of N8 mutant cells on inhibition of N8 eRNA transcription. Results of one representative experiment are presented ($n = 3$ biological repeats). **g**, NPC differentiation was confirmed on day 6 after differentiation by testing expression of ES cell-expressing genes and NPC-expressing genes. A representative experiment is presented ($n = 2$ biological repeats). In the box plots, line shows median and box shows 25th and 75th percentiles. *P* values calculated using Welch's two-sided *t*-test. Primer sets listed in Supplementary Table 2.
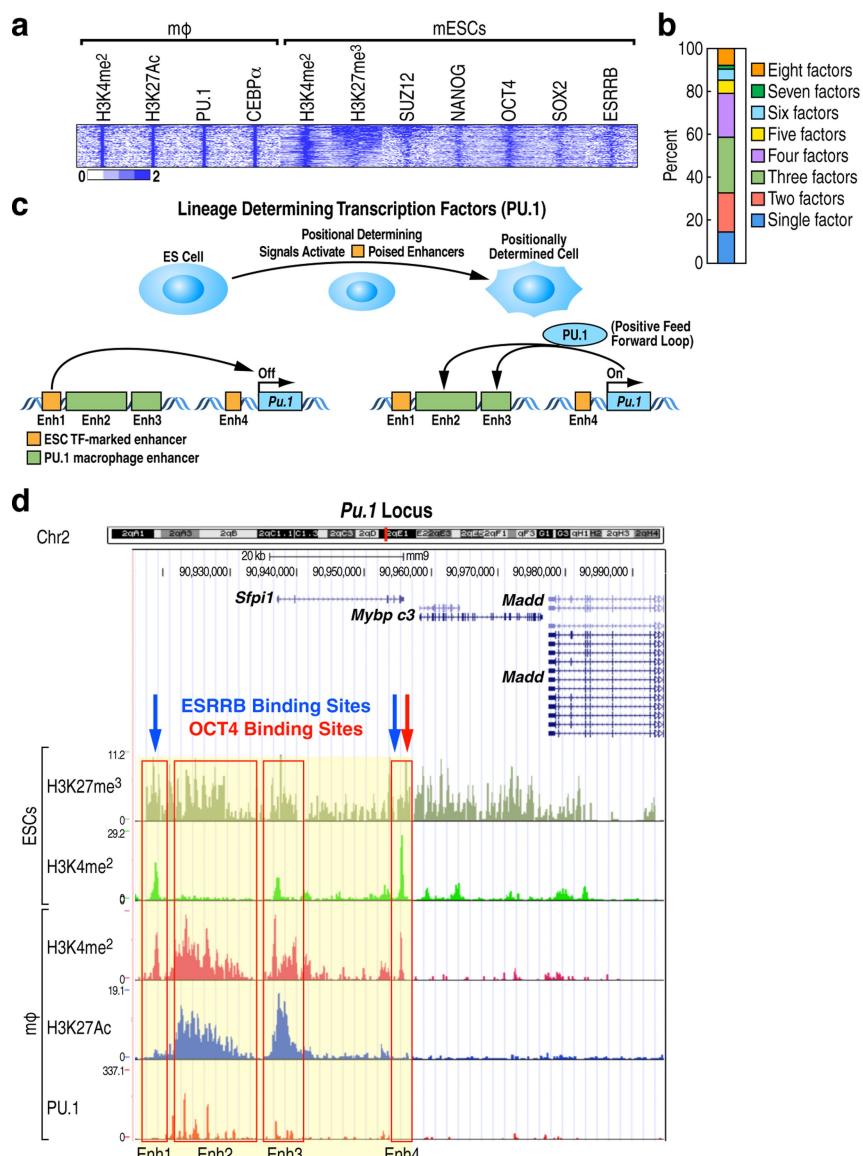
Extended Data Fig. 9 | Binding of ES cell transcription factors in poised enhancer. a, Profile of 214 macrophage enhancers with poised chromatin signatures in ES cells. **b**, ES cell transcription factors (OCT4, SOX2, NANOG, ESRRB, E2F1, TCFCP2L1, ZFX and KLF4) were analysed in 129 macrophage enhancers to test binding of ES cell transcription factors. These 129 macrophage enhancers were chosen according to the following criteria: high H3K27me3 in ES cells and high H3K27Ac (over 20 tags) in macrophages. Data from published sources are listed in Supplementary Table 1. **c**, Schematic diagram of *Pu.1* enhancers. **d**, Genome browser screenshot of the poised *Pu.1* enhancers. Blue arrow indicates ESRRB binding region and red arrow indicates OCT4 binding region. Red boxes indicate four clustered putative *Pu.1* enhancers.
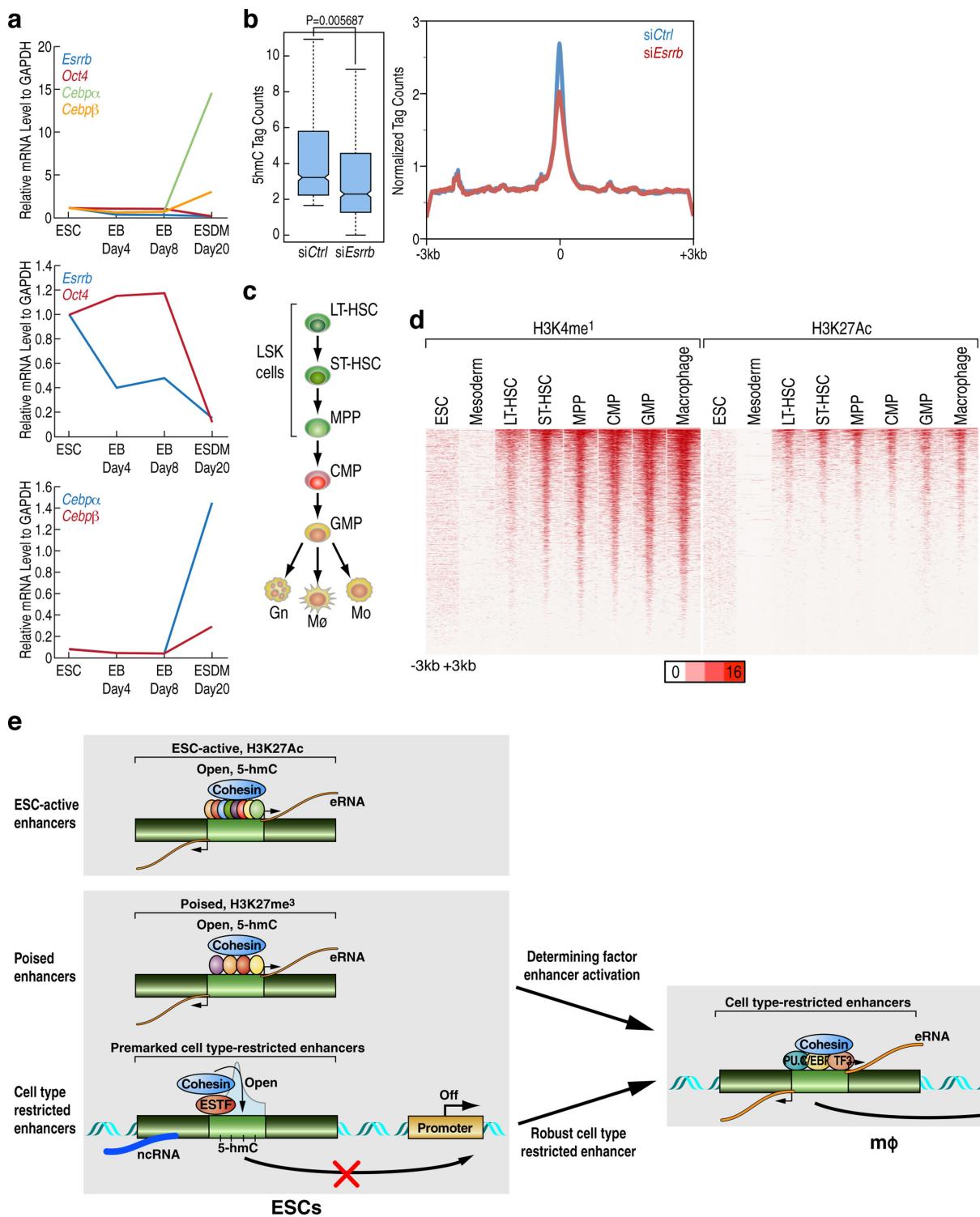
Extended Data Fig. 10 | Establishment and maintenance of 5-hmC on premarked enhancers and proposed model. a, *Esrrb*, *Oct4* and *Cebp* level measured at four different time points through differentiation process from ES cells to macrophages. **b**, 5-hmC level in 6,775 premarked enhancers in ES cells expressing scrambled (si*Ctrl*) or *Esrrb*-targeted (si*Esrrb*) siRNA. In the box plots, line shows median and box shows 25th and 75th percentiles. *P* values calculated using Welch's two-sided *t*-test. **c**, Schematic of the haematopoietic differentiation stages[24]. **d**, H3K4me1 and H3K27Ac in 6,775 premarked enhancers were analysed during haematopoietic differentiation (ES cells, mesoderm, LT-HSC, ST-HSC, MPP, CMP, GMP and macrophages), and presented

in heat map within a −3 kb/+3 kb window centred on PU.1. Data from published sources are listed in Supplementary Table 1. **e**, Model of 'premarked' lineage-determining and terminal-differentiation enhancer indicating that poised enhancers bind several ES cell transcription factors, whereas late-activated cell-specific enhancers are premarked by binding of a single ES cell transcription factor, causing chromatin opening, transcription of a non-coding RNA, and appearance of a 5-hmC mark in the area of the enhancer that may provide the molecular memory for what ultimately will be the PU.1–C/EBPα core from which eRNAs will be transcribed.

LETTER

Mechanism of NMDA receptor channel block by MK–801 and memantine

Xianqiang Song[1,7], Morten Ø. Jensen[2,7], Vishwanath Jogini[2,7], Richard A. Stein[3], Chia–Hsueh Lee[1,6], Hassane S. Mchaourab[3], David E. Shaw[2,4]* & Eric Gouaux[1,5]*

The NMDA (*N*-methyl-D-aspartate) receptor transduces the binding of glutamate and glycine, coupling it to the opening of a calcium-permeable ion channel[1]. Owing to the lack of high-resolution structural studies of the NMDA receptor, the mechanism by which ion-channel blockers occlude ion permeation is not well understood. Here we show that removal of the amino-terminal domains from the GluN1–GluN2B NMDA receptor yields a functional receptor and crystals with good diffraction properties, allowing us to map the binding site of the NMDA receptor blocker, MK-801. This crystal structure, together with long-timescale molecular dynamics simulations, shows how MK-801 and memantine (a drug approved for the treatment of Alzheimer's disease) bind within the vestibule of the ion channel, promote closure of the ion channel gate and lodge between the M3-helix-bundle crossing and the M2-pore loops, physically blocking ion permeation.

NMDA receptors comprise a heterotetrameric complex of two GluN1 subunits and two GluN2 subunits[2] and have a domain-layered architecture, with the amino-terminal domain (ATD) and the ligand- or agonist-binding domain (LBD) residing in the synaptic space and the transmembrane domain (TMD) spanning the membrane[3,4]. Activation of the ion channel requires binding of glutamate and glycine as well as voltage-dependent relief of magnesium block[5,6]; these cause membrane depolarization and calcium influx[7], both of which are critical for synaptic transmission and plasticity, and in cellular mechanisms of learning and memory[8]. Neurodegenerative disorders, chronic pain, stroke and schizophrenia have all been attributed to the dysfunction of NMDA receptors[9]. Over-activation of NMDA receptors is excitotoxic and contributes to neuronal damage after stroke or traumatic injury[10]. Furthermore, chronic NMDA receptor hyperactivity gives rise to the loss of neurons associated with Huntington's, Parkinson's and Alzheimer's diseases[11].

The ATDs are central to subunit-specific receptor assembly, to the probability, duration and deactivation rate of channel opening[12], and to the binding of allosteric modulators such as ifenprodil and Ro 25–698[13], which inhibit GluN1–GluN2B NMDA receptors and show promise for the treatment of Parkinson's and Alzheimer's diseases[14]. pH-dependent GluN2B-selective inhibitors are neuroprotective in traumatic brain injury, suggesting a strategy for treating brain injury with minimal side effects[15]. Moreover, small molecules that target the TMD, including MK-801 (dizocilpine) and memantine, show promise for the treatment of excitotoxicity-related disorders[16]. MK-801 is a neuroprotective agent in animal models of stroke, trauma and Parkinsonism but it can induce psychotic behaviour and neuronal degeneration[17]. The side effects of MK-801, which are probably a result of its high affinity for and long dwell time on the receptor, preclude its clinical application. By contrast, memantine, which binds more weakly to the ion channel, is well tolerated in clinical use. Memantine is used to treat Alzheimer's disease and early-onset epileptic encephalopathy[18,19]. Here we combine crystallography with long-timescale molecular dynamics simulations to

show how the GluN1–GluN2B NMDA receptor is blocked by MK-801 and memantine.

To facilitate structural analysis, we expressed GluN1–GluN2B NMDA receptor subunits[3] lacking the ATD, yielding the ΔATD receptor (Extended Data Fig. 1a, b), which was crystallized with glycine, glutamate and MK-801 (Fig. 1a, b). The ΔATD receptor crystallizes with two receptor complexes in the crystallographic asymmetric unit and, although each complex has a similar overall structure, there are differences in the relative positions of the LBD and TMD layers owing to the flexible nature of the LBD–TMD linkers (Fig. 1c–f, Extended Data Fig. 1c). The LBD layer adopts a 'dimer of dimers' organization, maintaining the same intra-dimer interface found in the intact GluN1–GluN2B NMDA receptor (the Δ2 receptor)[3]. In this ΔATD receptor structure, however, the LBD subunits have swapped partners relative to their arrangement in the Δ2 receptor (Fig. 1e, f), thus relaxing the GluN2B LBD–TMD linkers and precluding them from transmitting conformational changes of the LBD to the TMD[20] (Extended Data Fig. 2a–d, Supplementary Video 1). Double electron–electron resonance (DEER) experiments show that the LBD layer of the ΔATD receptor in solution adopts two arrangements of the LBD subunits, one that is consistent with the intact Δ2 NMDA receptor[3], and one that is consistent with that observed in the ΔATD crystal structure (Extended Data Fig. 2e). We hypothesize that, because a fraction of the ΔATD receptor molecules populate an intact receptor-like LBD layer arrangement, the ΔATD receptor retains agonist-induced ion channel gating. However, the ΔATD receptor also populates a conformation with a subunit-swapped LBD layer, a putatively inactive conformation that preferentially forms crystals. The propensity of the ΔATD receptor to adopt a LBD-swapped arrangement emphasizes the role of the ATD in defining the subunit arrangement of the LBD layer.

The ΔATD receptor retains agonist-induced ion channel activity, demonstrating that the ion channel is gated by glycine and glutamate (Extended Data Fig. 3a). Because the ΔATD receptor contains the thermostabilizing mutation G610R within the TMD, the dissociation constant (K_d) for MK-801 of $1.6 \pm 0.3\,\mu M$ is higher than that of the intact receptor, which has a K_d for MK-801 of $111.3 \pm 8.5\,nM$ (Fig. 2a). Indeed, when the G610R substitution is introduced into the Δ2 receptor, the K_d for MK-801 increases to $1.1 \pm 0.4\,\mu M$, similar to that of the ΔATD receptor (Fig. 2a). To ensure MK-801 occupancy within the ΔATD receptor TMD, we incubated the ΔATD receptor with a saturating concentration of MK-801 before crystallization. Diffraction-quality crystals emerged after two weeks and we solved the structure by molecular replacement (Extended Data Table 1).

The structure of the TMD of the ΔATD NMDA receptor exhibits a similar 'closed-blocked' state and pore radius as that of the Δ2 receptor[3] (Extended Data Fig. 3b, c). Moreover, the electron density of the ΔATD receptor TMD is superior to that of previously solved NMDA receptor structures, as evidenced by atomic *B* factors that are around 50% lower than those for the Δ2 receptor structure. As a result of this improved

[1]Vollum Institute, Oregon Health & Science University, Portland, OR, USA. [2]D. E. Shaw Research, New York, NY, USA. [3]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA. [4]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, USA. [5]Howard Hughes Medical Institute, Oregon Health & Science University, Portland, OR, USA. Present address: [6]Rockefeller University, New York, NY, USA. [7]These authors contributed equally: Xianqiang Song, Morten Ø. Jensen, Vishwanath Jogini.
*e-mail: David.Shaw@DEShawResearch.com; gouauxe@ohsu.edu
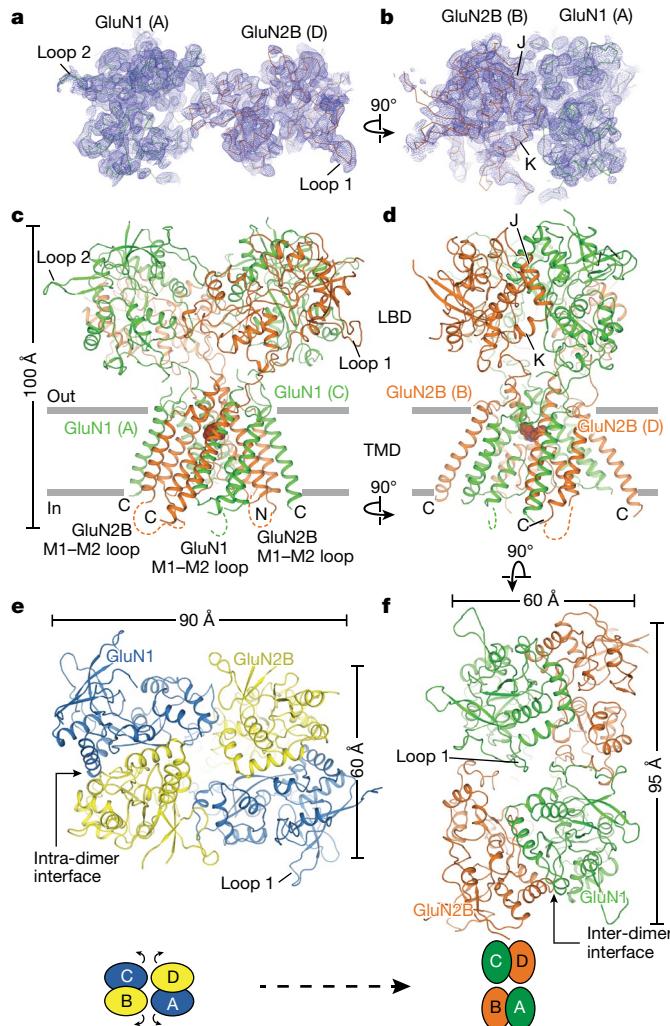
Fig. 1 | Architecture of the GluN1–GluN2B ΔATD NMDA receptor.
a, **b**, Composite omit maps (blue mesh) of GluN1 (green ribbon) and GluN2B (red ribbon) LBDs contoured at 1.0σ, showing the inter-dimer (**a**) and intra-dimer (**b**) interfaces. **c**, **d**, Side views of the ΔATD receptor, showing GluN1 (green) and GluN2B (orange) subunits. **e**, **f**, Top-down views of the Δ2 receptor (**e**) and the ΔATD receptor (**f**) from the extracellular side of the membrane.

Fig. 2 | MK-801 binding site defined by X-ray crystallography.
a, Saturation binding of [^{3}H]MK-801 to the intact Δ2 NMDA receptor and to the ΔATD receptor, which both feature the GluN1 G610R mutation, compared to the intact Δ2 G610 receptor, which harbours the native glycine residue at position 610. Error bars represent s.e.m. from triplicate measurements. **b**, Composite omit map (blue mesh) in the TMD region of the ΔATD GluN2B subunits (orange ribbon) and MK-801 (red stick), contoured at 1.0σ. **c**, **d**, Top-down (**c**) and side views (**d**) of MK-801 bound (carbon atoms in pink, nitrogen atom in blue) in the central vestibule, with the residues involved in the binding pocket shown in sticks and the α-carbons of the N612 residues (GluN2B) shown as spheres (**d**). **e**, Single-point binding of [^{3}H]MK-801 to the Δ2 G610 receptor and to indicated mutant receptors. Error bars represent s.e.m. from triplicate measurements. **f**, Top-down view of a slice through the TMD, where the protein surface is in solvent-accessible surface representation, showing the $F_o - F_c$ electron density (3.0σ, yellow mesh) of MK-801, the residues (sticks) involved in the binding pocket, and two tunnels (cyan) that connect the central vestibule to the cell membrane.

electron density, we are now able to define the binding site of MK-801 (Fig. 2b).

MK-801 resides within the channel vestibule and fits snugly in the binding pocket, with its nitrogen atom near N612 of GluN2B (Fig. 2c, d). The omit electron density map showed only electron density for the β-carbon of N612; we therefore did not build the complete side chain. Nevertheless, we modelled the side chain and determined that there is space for the complete side chain and that the oxygen of the primary amide can be positioned to interact with the amino group of MK-801. The two aromatic rings of MK-801 are juxtaposed next to the M3 helix of the GluN1 subunit, close to the V642 residues, whereas the methyl substituent is near L640 on the M3 helix of GluN2B (Fig. 2c, d). Accordingly, the substitutions V642A and V642L in GluN1, and L640A in GluN2B, extinguish MK-801 binding, consistent with previous studies[21] and with the importance of these residues for lining the binding pocket (Fig. 2e).

There are two 'tunnels' near the M2 helices of the TMD structure that connect the central vestibule to the interior of the membrane bilayer (Fig. 2f). The tunnels begin from the central vestibule and pass through cavities between the M2 and M3 helices of GluN1 and the adjacent M3 helices of GluN2B (Extended Data Fig. 4). The radii of the two tunnels

range from 1.3 to 2.9 Å and from 1.8 to 2.9 Å, respectively (Fig. 2f), large enough to accommodate the aliphatic tails of lipids. Indeed, in our molecular dynamics simulations, the lipid tails transiently entered these tunnels (Extended Data Fig. 4). We hypothesize that these tunnels could be transiently large enough to allow the diffusion of small molecules into the pore, perhaps explaining how small molecules are able to leave the closed, deactivated channel[22]. Nevertheless, our simulations find that both memantine and MK-801 enter the binding site of the open, activated, intact receptor via the aqueous path (Supplementary Videos 2–5).

The conformation of the ion channel region, consisting of the M2 helices, the pore loops and the M3 helices, differs between the GluN1 and GluN2B subunits (Fig. 3a). Residing on the tips of the pore loops, the long-studied N-site asparagines[23,24], N614 (GluN1) and N612 (GluN2B), together with the N + 1 site asparagines, N613 (GluN2B), have marked effects on channel block by MK-801[21] because they project their side chains into the vestibule and interact with MK-801 (Fig. 3b, c). These asparagines also contribute to the selectivity filter, modulating calcium permeability and voltage-dependent magnesium block[23,24]. The proximal pair of GluN2B subunits have M2 helices that

LETTER RESEARCH
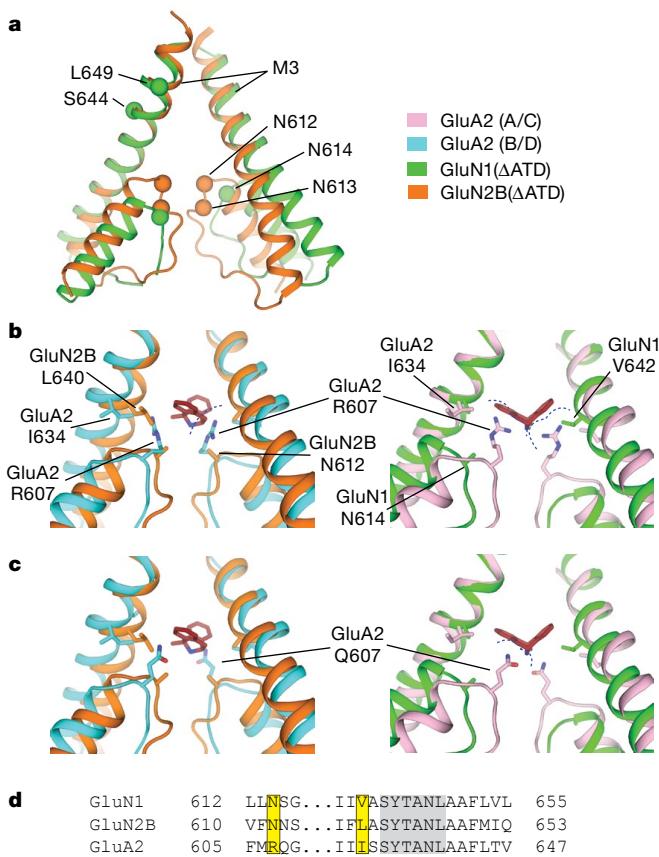
Fig. 3 | Steric clashes block MK-801 binding at AMPA receptors.
a, Superposition of the M2, pore loop and M3 elements of the GluN1 (green) and GluN2B (orange) subunits from the ΔATD receptor crystal structure. The α-carbons of key asparagine residues are shown as spheres. **b**, **c**, Superposition of elements of the ΔATD receptor from **a** with the equivalent elements of the GluA2 AMPA receptor (PDB code: 5VOT) with R607 (**b**) or Q607 (**c**). The NMDA receptor GluN2B subunits (orange) and GluN1 subunits (green) are superposed on the equivalent regions of the GluA2 AMPA receptor B and D subunits (cyan) or the A and C subunits (pink), respectively. All superpositions are based on the α-carbon atoms of the conserved SYTANL region. Dashed lines show likely steric clashes. **d**, Sequence alignment of the channel region between the NMDA receptor and AMPA receptor subunits. The residues involved in MK-801 binding and the corresponding GluA2 residues are highlighted in yellow. Residues of the SYTANL motif are highlighted in grey.

protrude more deeply into the ion channel vestibule in comparison to the GluN1 subunits, and therefore the GluN2B N + 1 site asparagine residues are located at the same level as the GluN1 N-site asparagines, and their side chains form a narrow constriction, consistent with previous studies[23] (Fig. 3a).

When the channels of the NMDA receptor and the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor are superposed, the AMPA receptor GluA2 R607 (the Q/R site) colocalizes with the critical N-site asparagines (N612) of GluN2B (Fig. 3b, c). As glutamine substitution in this position eliminates binding of MK-801[21], we hypothesize that the longer side chains of R607 or Q607 and the proximal-pore loop of the AMPA receptor A/C pair sterically hinder MK-801 binding (Fig. 3b, c). Sequence alignment also shows a disparity between the residues involved in MK-801 binding to the NMDA receptor and those at the equivalent positions in AMPA receptors (Fig. 3d). These residues are critical for MK-801 binding, as demonstrated by alanine substitutions or AMPA-like leucine substitutions, which eliminate binding of the blocker to the NMDA receptor (Fig. 2f). These structural and amino acid differences explain why MK-801 shows selectivity for NMDA receptors and why simple substitution of NMDA receptor residues into the corresponding positions

does not confer MK-801 binding and blocking of the ion channel to AMPA receptors.

To further characterize the position of MK-801, we used 3-iodo MK-801 (MK-801-I) (Fig. 4a), and showed that the halogenated analogue binds to the intact receptor with an inhibition constant (K_i) of 742 ± 1.3 nM (Fig. 4b). Despite extensive crystallization trials, we were unable to crystallize the MK-801-I–ΔATD-receptor complex. However, we were able to co-crystallize MK-801-I with the Δ2 NMDA receptor; subsequent anomalous diffraction studies allowed us to confirm that MK-801 binds in a similar position as in the ΔATD structure and to validate the placement of the ligand (Extended Data Table 1, data set 2; C2 space group). The anomalous difference Fourier maps show two prominent peaks in the channel vestibule (Fig. 4c, d), which are related by the two-fold axis along the ion channel pore and situated 'above' the pore loops of GluN1 and GluN2B subunits. This pose is consistent with the conclusion that MK-801 resides in the channel vestibule in two equivalent orientations, with position 3 of the MK-801 ring system on either side of the two-fold axis (Fig. 4c, d).

We performed atomic-level, long-timescale molecular dynamics simulations[25] to assess the binding position of MK-801 and its 3-iodo derivative. We began a simulation starting from the intact Δ2 NMDA receptor crystal structure inserted into a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC) membrane, with MK-801-I placed in the central vestibule at a position suggested by the anomalous difference densities and by the MK-801–ΔATD receptor complex structure (Fig. 4, simulation 1 in Supplementary Table 1). During this 30-μs molecular dynamics simulation, we observed two predominant poses of MK-801-I, in overall agreement with the crystallographic results (Fig. 4d). In both poses the amine group of the ligand points towards the selectivity filter, where it forms stable hydrogen bonds with the two pore-loop asparagine residues N614 (GluN1) and N612 (GluN2B) (Fig. 4d).

To gain insight into the binding of MK-801 to the ion channel without starting from a pose derived from the crystal structure, we performed long-timescale simulations of 'free' MK-801 binding to a model of the open, activated state of the receptor. We obtained a model of the open state by computationally introducing two mutations, A650R (GluN1) and A648R (GluN2B), within the nine-residue SYTANLAAF 'lurcher' motif[26]. These mutations are known to increase the open probability of the wild-type receptor[27], and led to the opening of the ion channel pore within a few microseconds (Fig. 4e, f, simulation 3), presumably owing to increased charge repulsion at the M3 bundle-crossing region. After observing that the resulting open, but still non-permeating, mutant channel maintained hydration of its pore cavity on a microsecond timescale, we reversed the mutations and verified that the wild-type channel remained in a stable open and now conducting state, with an average conductance of 4.8 ± 1.4 pS (simulations 4–17). The experimental permeation rate is 51.4 ± 2.4 pS[28]; this disagreement is unsurprising, as it is known that molecular dynamics simulations typically do not accurately reproduce experimental ion permeation rates[29], owing to force field shortcomings.

Subsequently, we introduced MK-801 into the aqueous phase of the system and performed simulations of MK-801 binding to the intact, open receptor (simulations 18–21). Binding of MK-801 to the pore vestibule occurred along the ion permeation pathway and with a mean binding time of 0.78 ± 0.10 μs (simulations 20 and 21 were conducted at zero transmembrane voltage; application of voltage in simulations 18 and 19 did not significantly decrease the binding time, as shown in Extended Data Fig. 5). The binding event was followed by closure of the pore at the bundle-crossing region, which is likely to be driven by a hydrophobic collapse of residues around the nonpolar part of the blocker, thus giving rise to a closed-blocked state of the receptor, which can be deduced from the root mean square deviation (r.m.s.d.) of the M3 bundle-crossing region with respect to the closed-state Δ2 crystal structure (Extended Data Fig. 6a, b). The r.m.s.d. value increases in proportion to the applied voltage when the pore assumes its open, non-permeating and open, permeating states, whereas upon MK-801

© 2018 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

NATURE | www.nature.com/nature
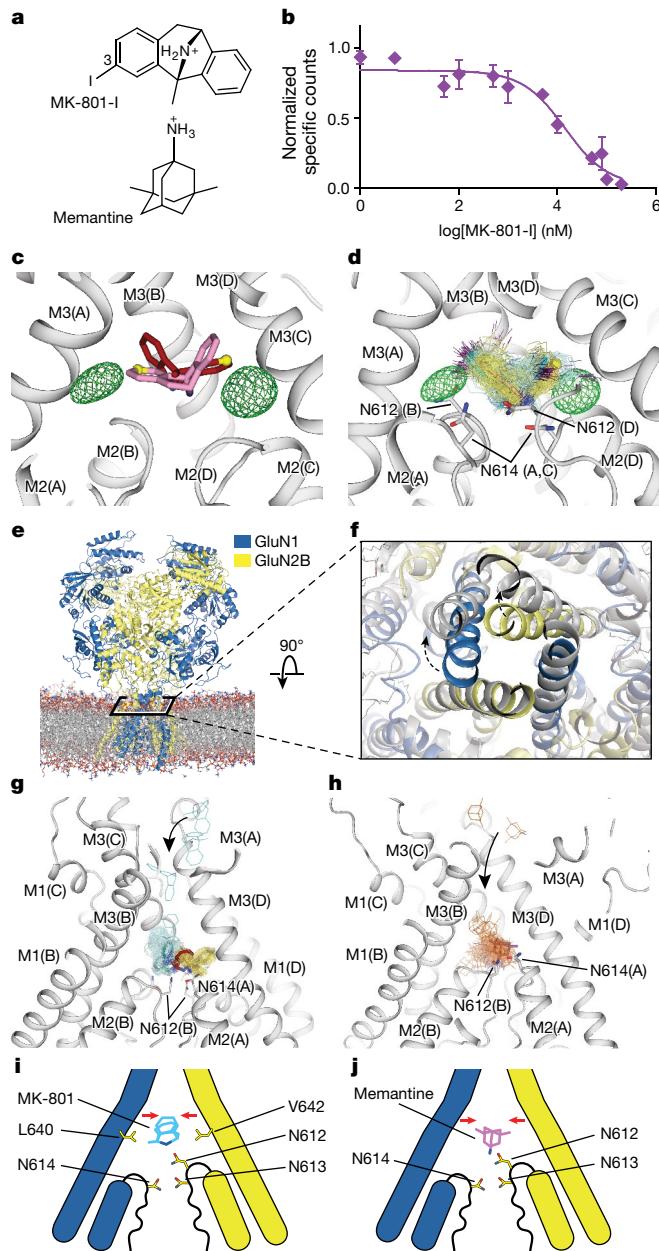
Fig. 4 | Mechanism of MK-801 and memantine binding. a, The chemical structures of MK-801-I and memantine. **b**, Competition binding plots of 'cold' MK-801-I to the [^{3}H]MK-801-bound Δ2 G610 receptor. The plot shows data from a representative experiment with error bars representing s.e.m. from triplicate measurements. **c**, The iodine anomalous density (green mesh) shown together with the positions of the MK-801 molecule in the crystal structure (red sticks) and the pseudo-two-fold related binding pose (pink sticks). The carbon atoms at the 3 position are shown as yellow spheres. **d**, Positions of MK-801-I obtained from molecular dynamics simulations (simulation 1) initiated from MK-801-I docked in the binding pocket of the Δ2 structure. The movements of the iodine atoms (purple) of MK-801-I are clustered into two populations, in agreement with the anomalous difference densities (green mesh). In each of its two predominant poses (cyan and yellow sticks), MK-801-I forms hydrogen bonds with N614 (GluN1) and N612 (GluN2B) (grey sticks). **e**, Simulation snapshot of Δ2 NMDA receptor (GluN1 in blue and GluN2B in yellow) embedded in POPC lipid membrane (red and grey lines). **f**, The open pore (grey) obtained in simulation 3, superposed onto the closed pore of the Δ2 receptor crystal structure. Arrows indicate the transition that occurs in gate opening. **g, h**, Snapshots of MK-801 (**g**; cyan, simulation 20) and memantine (**h**; orange, simulation 24) from free-binding simulations in which they bind to the open state of the Δ2 receptor (at zero transmembrane voltage). The blockers enter the pore by the aqueous path (black arrows). MK-801 exhibits two distributions of binding poses (cyan and yellow sticks), overlapping with MK-801 in the crystal structure (red stick), but memantine shows a predominant pose (purple). Both MK-801 and memantine interact with asparagine residues (grey sticks) on the pore loops. **i, j**, Schematic representations of MK-801 and memantine binding sites, respectively. Both channel blockers induce channel closure (red arrows) while blocking the pore and adopting similar interactions with key asparagine residues.

binding, the r.m.s.d. value drops, approaching the value obtained in our simulations of the closed-apo pore (simulation 2). Over the last part of each binding simulation (for example, the last 40 μs of simulation time in simulation 20, a 60-μs simulation), MK-801 predominantly assumed two distinct binding poses, similar to those found for MK-801-I in simulation 1, where the amine group of MK-801 formed stable hydrogen bonds with the two pore-loop asparagine residues (Fig. 4g, i).

Upon comparison of the simulated MK-801-bound receptor with the experimental structure, we noticed that MK-801 in the crystal structure resides in an intermediate pose between the two predominant binding poses observed in the simulations, indicating that the electron density of MK-801 in the crystal structure is an average of two distinct poses that are related by two-fold symmetry (Fig. 4g, i). Additional simulations with memantine (Fig. 4h, j, simulations 24–27), indicate that this compound binds at a location nearly identical to the binding location of MK-801. Rapid entry (mean binding time 0.14 ± 0.02 μs) was again followed by a hydrophobic collapse of residues around the blocker, giving rise to a closed–blocked state of the receptor (Extended Data Figs. 5b, 6c). Memantine, however, predominantly assumes a single binding pose, perhaps owing to its pseudo-symmetrical structure (Fig. 4h, j). We estimated the absolute free energy of binding to the Δ2 receptor,

and found $K_d \approx 7.6$ μM for memantine and $K_d \approx 0.08$ μM for MK-801. This 100-fold difference in relative affinity is similar to that between the experimental affinities, 147.4 μM and 1.1 μM (Extended Data Fig. 7a), noting that the errors in the free energies are substantial (Extended Data Fig. 7). We speculate that the off-rate of memantine is approximately 100-fold faster than that of MK-801, because the on-rates and the experimental and computational binding constants are similar for both ligands. The faster off-rate of memantine is important for its clinical efficacy[30]. Both MK-801 and memantine readily entered and bound to the pore cavity regardless of voltage, but an applied voltage drove the blockers deeper into the selectivity filter, where they formed hydrogen bonds with the N + 1 asparagine residue (N613 of GluN2B) (Extended Data Fig. 8). This might explain the experimental observation that the presence of the N + 1 asparagine residue leads to voltage dependency of MK-801 and memantine binding.

We observed large fluctuations in the selectivity filter in these long-timescale simulations, both because of force-field deficiencies and, potentially, because our simulations begin with structures for which the experimental information in the selectivity filter is incomplete. We therefore investigated whether restraining the selectivity filter close to the crystal structure would affect the binding poses of MK-801 and memantine (Extended Data Fig. 9). We conclude that conformational distributions of the selectivity filter do not affect the overall pose or position of memantine and that it blocks the pore in predominantly one pose, whereas MK-801 blocks the pore in two symmetry-related poses, with the same key interactions between the selectivity filter and each pore blocker. Additional control simulations show that MK-801 also adopts these two symmetry-related binding poses in the absence of torsional corrections to the selectivity filter (see Methods and Extended Data Fig. 9.)

The structure of the ΔATD GluN1–GluN2B receptor shows how the absence of the ATDs allows the LBD layer to adopt a previously unseen conformation (Supplementary Video 1). This deletion enables formation of crystals that diffract to sufficiently high resolution to allow us to position the ion channel blocker MK-801 . MK-801 and memantine bind within the TMD vestibule and block ion conduction by physical occlusion of the permeation pathway and by promoting closure of the

ion channel gate at the M3-helix-bundle crossing, yielding a closed-blocked state. MK-801 binds in two, two-fold-related poses whereas memantine binds in a single predominant pose. Our studies define the molecular basis for ion channel block by MK-801 and memantine and lay foundations for the development of small molecules with desirable interactions with the NMDA receptor, such as receptor-subunit selectivity and affinity.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0039-9.

1. Traynelis, S. F. et al. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacol. Rev.* **62**, 405–496 (2010).
2. Paoletti, P. Molecular basis of NMDA receptor functional diversity. *Eur. J. Neurosci.* **33**, 1351–1365 (2011).
3. Lee, C. H. et al. NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature* **511**, 191–197 (2014).
4. Karakas, E. & Furukawa, H. Crystal structure of a heteromeric NMDA receptor ion channel. *Science* **344**, 992–997 (2014).
5. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. Voltage-dependent block by Mg^{2+} of NMDA responses in spinal cord neurones. *Nature* **309**, 261–263 (1984).
6. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. Magnesium gates glutamate-activated channels in mouse central neurones. *Nature* **307**, 462–465 (1984).
7. Johnson, J. W. & Ascher, P. Glycine potentiates the NMDA response in cultured mouse brain neurons. *Nature* **325**, 529–531 (1987).
8. Bliss, T. V. P. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
9. Paoletti, P., Bellone, C. & Zhou, Q. NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nat. Rev. Neurosci.* **14**, 383–400 (2013).
10. Simon, R. P., Swan, J. H., Griffiths, T. & Meldrum, B. S. Blockade of N-methyl-D-aspartate receptors may protect against ischemic damage in the brain. *Science* **226**, 850–852 (1984).
11. Parsons, M. P. & Raymond, L. A. Extrasynaptic NMDA receptor involvement in central nervous system disorders. *Neuron* **82**, 279–293 (2014).
12. Yuan, H., Hansen, K. B., Vance, K. M., Ogden, K. K. & Traynelis, S. F. Control of NMDA receptor function by the NR2 subunit amino-terminal domain. *J. Neurosci.* **29**, 12045–12058 (2009).
13. Karakas, E., Simorowski, N. & Furukawa, H. Subunit arrangement and phenylethanolamine binding in GluN1/GluN2 NMDA receptors. *Nature* **475**, 249–253 (2011).
14. Hu, N. W., Klyubin, I., Anwyl, R. & Rowan, M. J. GluN2B subunit-containing NMDA receptor antagonists prevent Aβ-mediated synaptic plasticity disruption in vivo. *Proc. Natl Acad. Sci. USA* **106**, 20504–20509 (2009).
15. Yuan, H. et al. Context-dependent GluN2B-selective inhibitors of NMDA receptor function are neuroprotective with minimal side effects. *Neuron* **85**, 1305–1318 (2015).
16. Parsons, C. G. et al. Comparison of the potency, kinetics and voltage-dependency of a series of uncompetitive NMDA receptor antagonists in vitro with anticonvulsive and motor impairment activity in vivo. *Neuropharmacology* **34**, 1239–1258 (1995).
17. Kovacic, P. & Somanathan, R. Clinical physiology and mechanism of dizocilpine (MK-801): electron transfer, radicals, redox metabolites and bioactivity. *Oxid. Med. Cell. Longev.* **3**, 13–22 (2010).
18. Reisberg, B. et al. Memantine in moderate-to-severe Alzheimer's disease. *N. Engl. J. Med.* **348**, 1333–1341 (2003).
19. Pierson, T. M. et al. GRIN2A mutation and early-onset epileptic encephalopathy: personalized therapy with memantine. *Ann. Clin. Transl. Neurol.* **1**, 190–198 (2014).
20. Wollmuth, L. P. & Sobolevsky, A. I. Structure and gating of the glutamate receptor ion channel. *Trends Neurosci.* **27**, 321–328 (2004).
21. Kashiwagi, K. et al. Channel blockers acting at N-methyl-D-aspartate receptors: differential effects of mutations in the vestibule and ion channel pore. *Mol. Pharmacol.* **61**, 533–545 (2002).
22. Sobolevsky, A. I., Koshelev, S. G. & Khodorov, B. I. Interaction of memantine and amantadine with agonist-unbound NMDA-receptor channels in acutely isolated rat hippocampal neurons. *J. Physiol. (Lond.)* **512**, 47–60 (1998).
23. Wollmuth, L. P., Kuner, T., Seeburg, P. H. & Sakmann, B. Differential contribution of the NR1- and NR2A-subunits to the selectivity filter of recombinant NMDA receptor channels. *J. Physiol. (Lond.)* **491**, 779–797 (1996).
24. Wollmuth, L. P., Kuner, T. & Sakmann, B. Intracellular Mg^{2+} interacts with structural determinants of the narrow constriction contributed by the NR1-subunit in the NMDA receptor channel. *J. Physiol. (Lond.)* **506**, 33–52 (1998).
25. Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O. & Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* **19**, 120–127 (2009).
26. Chang, H. R. & Kuo, C. C. The activation gate and gating mechanism of the NMDA receptor. *J. Neurosci.* **28**, 1546–1556 (2008).
27. Murthy, S. E., Shogan, T., Page, J. C., Kasperek, E. M. & Popescu, G. K. Probing the activation sequence of NMDA receptors with lurcher mutations. *J. Gen. Physiol.* **140**, 267–277 (2012).
28. Stern, P., Cik, M., Colquhoun, D. & Stephenson, F. A. Single channel properties of cloned NMDA receptors in a human cell line: comparison with results from *Xenopus* oocytes. *J. Physiol. (Lond.)* **476**, 391–397 (1994).
29. Jensen, M. Ø., Jogini, V., Eastwood, M. P. & Shaw, D. E. Atomic-level simulation of current-voltage relationships in single-file ion channels. *J. Gen. Physiol.* **141**, 619–632 (2013).
30. Lipton, S. A. Paradigm shift in neuroprotection by NMDA receptor blockade: memantine and beyond. *Nat. Rev. Drug Discov.* **5**, 160–170 (2006).

Author contributions X.S. and E.G. designed the project; X.S. and C.-H.L. developed the constructs for crystallization; X.S. performed protein purification, crystallography, electrophysiology and biochemical analysis; M.Ø.J. and V.J. performed and analysed, and D.E.S. participated in the oversight of, the molecular dynamics simulations; R.A.S. and H.S.M. carried out the DEER experiments; and X.S. and E.G. wrote the manuscript with contributions from all authors.

Competing interests The authors declare no competing interests.

METHODS

Receptor constructs. The construct of the ΔATD NMDA receptor is derived from the previously reported *Xenopus laevis* GluN1–GluN2B NMDA receptor Δ2 construct[3] and cloned into pEG BacMam[31,32] for virus-mediated expression in suspension-adapted cells. Residues 23–393 were removed from GluN1, residues 1–399 were removed from GluN2B, and the human placental alkaline phosphatase signal sequence (MLGPCMLLLLLLLGLRLQLSLG) was added to the N terminus of the ΔATD GluN2B subunit[33]. The Δ2 construct was used for co-crystallization with iodo-MK-801 (Sanofi) for the anomalous diffraction studies. The DEER construct was designed with alanine substitutions of the endogenous cysteines, GluN1(C457A) and GluN2B(C458A), and a single cysteine GluN2B(R739C) substitution.

Expression and purification. HEK293S GnTI⁻ cells[34] were transduced using P2 or P3 BacMam virus (titre $> 3 \times 10^6$ plaque-forming units (p.f.u.)) when the cell density reached 3×10^6 cells/ml at a multiplicity of infection (MOI) of 1:1 (GluN1:GluN2B) and incubated at 37 °C. At 14 h post transduction, 10 mM sodium butyrate was added to the cultures and the cultures were incubated at 30 °C for another 46 h.

Cells were harvested and collected by centrifugation at 6,200*g* for 20 min. Cells were disrupted by sonication (4 s on, 8 s off for 10 min, power level 6.5) in 150 mM NaCl, 20 mM Tris-HCl pH 8.0, at a v/v ratio of 25 ml buffer per litre culture. The cell debris was removed by centrifugation at 6,000*g* for 20 min and the membrane was collected by ultracentrifugation at 125,000*g* for 1 h, at 4 °C. Membranes were homogenized in 150 mM NaCl, 20 mM Tris-HCl pH 8.0 and solubilized in a buffer containing 1% lauryl maltose neopentyl glycol (MNG-3; Anatrace NG310), 0.8 mM aprotinin, 2 mg/ml leupeptin, 2 mM pepstatin A, 1 mM phenylmethylsulfonyl fluoride, 1 mM glutamate, 1 mM glycine and 2 mM cholesteryl hemisuccinate (CHS, Anatrace CH210) for 2 h, with stirring at 4 °C. Insoluble material was removed by centrifugation at 125,000*g*. The supernatant was loaded onto a Strep-Tactin column (∼1 ml resin per litre culture, IBA 2-4010-025) and eluted with buffer containing 5 mM desthiobiotin. The receptor was concentrated and the GFP, octa-histidine and StrepII tags were removed by treatment with 3 C protease at 1:20 ratio (w:w) in tandem with endoglycosidase H treatment for 8 h at 4 °C before size-exclusion chromatography (SEC). The concentrated GluN1–GluN2B receptor was loaded onto a Superose 6 16/300 column equilibrated with 400 mM NaCl, 20 mM MES pH 6.5, 1 mM dodecyl maltoside (C12M) and 0.2 mM CHS. Peak fractions were pooled and concentrated to 2 mg/ml.

Crystallization and cryoprotection. Prior to crystallization, 10 mM 6-cyclohexyl-hexyl-β-ᴅ-maltoside (CYMAL-6, Anatrace C326LA), 10 mM glutamate, 10 mM glycine and 1 mM MK-801 were added to the GluN1–GluN2B protein, incubated on ice for 14–16 h and the protein solution was then centrifuged at 70,000*g* for 40 min at 4 °C. Crystals were grown by vapour diffusion using a reservoir solution composed of 100 mM MES pH 6.5, 200 mM NaF and 19–20% polyethylene glycol 3350. Crystals were cryoprotected with mother liquor supplemented with 14% ethylene glycol. To obtain iodo-MK-801-complexed crystals, 1 mM iodo-MK-801 was co-crystallized with the protein in 18–20% pentaerythritol ethoxylate, 100 mM NaCl, 100 mM MgCl₂ and 100 mM MES pH 6.5. Crystals were cryoprotected with crystallization buffer supplemented with 12% ethylene glycol. All crystals were obtained by hanging drop vapour diffusion with a drop ratio of 1 or 2 μl protein to 1 μl reservoir solution at 20 °C.

Structure determination and analysis. X-ray diffraction data sets were collected at the Advanced Light Source on beamlines 8.2.1 and 5.0.2 and the Advanced Photon Source (APS) on beamlines 24ID-E and 24ID-C. Diffraction images were indexed, integrated and scaled by XDS[35] or HKL2000, with the microdiffraction assembly method[36]. The best data set was collected at APS 24ID-E and the data extended to Bragg spacings at 3.9 Å, 3.7 Å and 3.4 Å resolution along a^*, b^* and c^*[37]. The structure was determined by molecular replacement with Phaser[38]. Model building and crystallographic refinement were carried out using the computer graphics program Coot[39] and the crystallographic refinement software package Phenix[40]. We were guided in building the ΔATD TMD by referring to the intact receptor structure[3], using aromatic side chains to define sequence register. The model was refined to a nominal resolution of 3.6 Å with reasonable *R* factors. Stereochemistry of the model was evaluated by MolProbity[41] and figures were created using Pymol. In analysing the tunnels of the transmembrane domain, because multiple side chain groups were truncated upon model building owing to insufficient electron density, we incorporated the side chains in order to analyse the path and radii of the tunnels. The side chains were built using SWISS-MODEL[42] and the tunnel size and radius were measured using CAVER[43].

Two-electrode voltage clamp. Recombinant NMDARs were expressed in *Xenopus laevis* oocytes after cytoplasmic injection of a 50-nl mixture of RNAs encoding the particular GluN1 and GluN2B subunits (at 500 ng/μl concentration, ratio 1:1) and stored at 16 °C in Barth's solution (in mM: 10 HEPES, pH 7.5, 88 NaCl, 2.4 NaHCO₃, 1 KCl, 0.33 Ca(NO₃)₂, 0.91 CaCl₂, 0.82 MgSO₄) supplemented with 50 μM AP5 and 100 μg/ml gentamicin. Recordings were performed using a bath solution containing (in mM) 5 HEPES, pH 7.3, 100 NaCl, 2.8 KCl and 0.3 BaCl₂. EDTA (10 μM final concentration) was added to all bath solutions. Currents were elicited by simultaneous application of 300 μM glycine and 300 μM glutamate. The holding potential was −60 mV.

Ligand-binding assays. Binding constants were measured by the scintillation proximity assay (SPA)[44]. SPAs were set up in triplicate at a final volume of 100 μl in SPA buffer (20 mM Tris pH 8.0, 150 mM NaCl, 0.01% MNG-3 and 0.02 mM CHS). NMDA receptor (50 nM, purified from affinity chromatography) was incubated with 0.5 mg/ml Ysi-Cu SPA beads, [³H]MK-801 and 1 mM each of glutamate and glycine. Non-specific binding was determined by the addition of 400 μM phencyclidine (PCP). Inhibition constants were determined by the SPA using 50 nM receptor, 0.5 mg/ml Ysi-Cu SPA beads, 2 μM [³H]MK-801 and 1 mM each of glutamate and glycine and varying concentrations of 'cold' iodo-MK-801. The samples were gently mixed and incubated at room temperature for 2 h before the counts were measured. The counts were analysed using GraphPad Prism.

Molecular dynamics simulations. All molecular dynamics simulations were based on the crystal structure of the Δ2 GluN1–GluN2B intact, heterotetrameric receptor (PDB code: 4TLM; GluN1: chains A and C; GluN2B: chains B and D). Missing loops (chain A: 547–549, 583–590, 650–653; chain B: 532–540, 576–584, 638–649, 789–797; chain C: 482–502, 540–550, 583–591, 786–796; chain D: 383–396, 434–445, 534–540, 789–793) were built using MOE[45]. Protonation states of all residues were chosen to correspond to a pH value of 7. The receptor was embedded in a POPC lipid bilayer hydrated by 0.2 M NaCl. The system contained ∼507,000 atoms and measured ∼$160 \times 160 \times 188$ Å³. To impose a transmembrane voltage difference, *V*, we applied constant electric fields of −$0.0125 \le E \le -0.075$ kcal mol⁻¹ Å⁻¹ e⁻¹), corresponding to $-100 \le V \le -600$ mV[29,46]. To study the binding of the endogenous ligand analogues 1AC (1-aminocyclopropane-1-carboxylic acid) and JEG (trans-1-aminocyclobutane-1,3-dicarboxylic acid), we simulated the LBD of the GluN1 subunit in isolation (simulations 34 and 35) and the LBD dimer (LBD1: residues 397–544, 661–789; LBD2: residues 399–534 and 647–785; simulations 36–40); both constructs were derived from simulation 2 of the intact heterotetrameric receptor. All simulations were performed in the NPT ensemble at 310 K using Anton[47].

Simulations of receptor-bound, MK-801-I (simulation 1) were initiated with the ligand docked into the vestibule such that the ligand iodine atom initially occupied one of the two experimentally obtained anomalous difference densities (Fig. 4). A control simulation (simulation 2) of the same system in the absence of MK-801 was performed for comparison and showed no notable differences with simulation 1 in either the pore vestibule or elsewhere in the receptor.

To simulate free binding of MK-801 to a model of the open-state receptor, we first performed a simulation (simulation 3) to open the closed pore by introducing two mutations, A650R (GluN1) and A648R (GluN2B), near the M3-helix-bundle-crossing region, within the nine-residue SYTANLAAF lurcher motif[26]. Such mutations are known to enhance the constitutive activity of the receptor[27]. Next we performed a series of simulations (simulations 4–17) in which we reversed the mutations and verified that the wild-type channel remained in a stable open and now conducting state. All of these simulations were performed at negative transmembrane voltage. Subsequently, simulations of free binding of MK-801 to the open pore were carried out on the back-mutated wild-type receptor with and without a negative voltage (simulations 18–23). To increase the ligand-binding on-rate, we imposed a spherical flat-bottom harmonic restraint between MK-801 and the receptor vestibule, centred at the midpoint of 642 (GluN1) and 640 (GluN2B) with a 30.0 Å radius. Finally, we performed a series of comparative simulations of free binding of memantine, a known NMDA receptor antagonist closely related to MK-801 (simulations 24–32).

In all simulations, harmonic restraints were imposed on the endogenous ligand analogues (4-[(1*R*,2*S*)-3-(4-benzylpiperidin-1-yl)-1-hydroxy-2-methylpropyl]phenol (QEM), trans-1-aminocyclobutane-1,3-dicarboxylic acid (JEG) and 1-aminocyclopropanecarboxylic acid (1AC), which are present in the crystal structure) in order to prevent unbinding from their respective binding sites within the extracellular LBD. In separate simulations of the intact receptor (simulation 33), a single LBD (simulations 34 and 35) and the LBD dimer (simulations 36 and 40), all without ligand restraints, the two charged ligands (JEG and 1AC) consistently unbound too rapidly, with mean unbinding times of 0.4 ± 0.1 μs (JEG, $n = 5$) and 5.1 ± 1.9 μs (1AC, $n = 7$). The uncharged QEM did not unbind. An additional control simulation (simulation 41) of a crystal structure of the GluN1–GluN2A LBD dimer, co-crystallized with glutamate and glycine (PDB code: 5H8F[48]), consistently resulted in excessively rapid unbinding of the charged ligand (glutamate).

Backbone torsional corrections to residues 614–618 (chains A and C) and 612–616 (chains B and D) (which lowered the backbone torsional potential by 4.6 kcal mol⁻¹ at the crystallographic φ and ψ values, θ_0, relative to the value of the potential at $\theta_0 + 180°$) were applied to ensure that the conformation of the selectivity filter remained intact during the simulations. We imposed these

corrections because the free-energy minimum in the force field at the (correct) crystal structure is unfortunately too shallow to ensure structural integrity of the filter region on the simulation timescales reported here[29]. In a subset of the simulations, additional r.m.s.d. restraints were imposed on the selectivity filter residues to study the effect of restraints on the binding poses. We found that the resulting conformational distributions of the selectivity filter influenced some details of the pore-blocker binding poses. We also performed a set of 13 control simulations (simulations 42–54; lengths ranging from 10–30 μs) of the full-length Δ2 receptor without backbone torsional corrections and analysed the binding poses of MK-801-I and MK-801 before deterioration of the filter region. We found that the resulting binding poses were similar to those obtained in the simulations with the filter corrections (Extended Data Fig. 9).

We used the CHARMM27 force field for the protein, ions and water[49,50], and the CHARMM36 force field for POPC[51]. Modified partial charges were used for Glu, Asp and Arg residues[52]. We used the CHARMM general force field[53,54] for all ligands and reparameterized all partial charges and torsional angles.

The aggregate simulation time was ~900 μs and individual simulations ranged from 12 μs to 60 μs. Additional simulation details are given in Supplementary Table 1.

Spin labelling and DEER experiments. The DEER construct was expressed in HEK293S GnTI⁻ cells using the same protocol as for the ΔATD receptor. The Strep-Tactin-purified GFP-fused protein was digested with 3 C protease at 4 °C when dialysed overnight against labelling buffer (150 mM NaCl, 20 mM Tris-HCl pH 7.2, 0.1% MNG-3 and 0.2 mM CHS). The labelling reaction was started by adding a tenfold molar excess of 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl-methanethiosulfonate (MTSSL) and incubating the mixture at room temperature in the dark. After 1 h, another tenfold MTSSL was added and incubated for an additional 1 h. The protein was kept on ice overnight, followed by gel filtration the next day. Peak fractions were pooled and concentrated. Electron paramagnetic resonance (EPR) spectra were obtained using continuous wave EPR as described previously[55]. DEER experiments were performed using a standard four-pulse protocol[56] and the resulting signals were analysed assuming that the distance distributions, $P(r)$, consist of a sum of Gaussians[57].

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. The coordinates for the ΔATD NMDA receptor structure have been deposited in the Protein Data Bank under the accession code 5UN1. The simulation results are available from the corresponding authors upon request.

31. Dukkipati, A., Park, H. H., Waghray, D., Fischer, S. & Garcia, K. C. BacMam system for high-level expression of recombinant soluble and membrane glycoproteins for structural studies. *Protein Expr. Purif.* **62**, 160–170 (2008).
32. Baconguis, I. & Gouaux, E. Structural plasticity and dynamic selectivity of acid-sensing ion channel–spider toxin complexes. *Nature* **489**, 400–405 (2012).
33. Inouye, H., Barnes, W. & Beckwith, J. Signal sequence of alkaline phosphatase of *Escherichia coli*. *J. Bacteriol.* **149**, 434–439 (1982).
34. Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* **99**, 13419–13424 (2002).
35. Kabsch, W. XDS. Acta Crystallogr. D Biol. *Crystallogr.* **66**, 125–132 (2010).
36. Hanson, M. A. et al. Crystal structure of a lipid G protein-coupled receptor. *Science* **335**, 851–855 (2012).
37. Armstrong, N., Jasti, J., Beich-Frandsen, M. & Gouaux, E. Measurement of conformational changes accompanying desensitization in an ionotropic glutamate receptor. *Cell* **127**, 85–97 (2006).
38. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 32–41 (2007).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
40. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, (1948–1954 (2002).
41. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
42. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure and homology modelling. *Bioinformatics* **22**, 195–201 (2006).
43. Chovancova, E. et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLOS Comput. Biol.* **8**, e1002708 (2012).
44. Hart, H. E. & Greenwald, E. B. Scintillation proximity assay (SPA)—a new method of immunoassay. Direct and inhibition mode detection with human albumin and rabbit antihuman albumin. *Mol. Immunol.* **16**, 265–267 (1979).
45. Vilar, S., Cozza, G. & Moro, S. Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. *Curr. Top. Med. Chem.* **8**, 1555–1572 (2008).
46. Roux, B. The membrane potential and its representation by a constant electric field in computer simulations. *Biophys. J.* **95**, 4205–4216 (2008).
47. Shaw, D. E. et al. Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In *Proc. International Conference for High Performance Computing, Networking, Storage and Analysis* 41–53 https://doi.org/10.1109/SC.2014.9 (2014).
48. Hackos, D. H. et al. Positive allosteric modulators of GluN2A-containing NMDARs with distinct modes of action and impacts on circuit function. *Neuron* **89**, 983–999 (2016).
49. MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
50. Mackerell, A. D. Jr, Feig, M. & Brooks, C. L. III Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**, 1400–1415 (2004).
51. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
52. Jensen, M. Ø. et al. Mechanism of voltage gating in potassium channels. *Science* **336**, 229–233 (2012).
53. Vanommeslaeghe, K. et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690, (2010).
54. Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D. Jr. Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **33**, 2451–2468 (2012).
55. Mishra, S. et al. Conformational dynamics of the nucleotide binding domains and the power stroke of a heterodimeric ABC transporter. *eLife* **3**, e02740 (2014).
56. Jeschke, G., Koch, A., Jonas, U. & Godt, A. Direct conversion of EPR dipolar time evolution data to distance distributions. *J. Magn. Reson.* **155**, 72–82 (2002).
57. Stein, R. A., Beth, A. H. & Hustedt, E. J. A straightforward approach to the analysis of double electron-electron resonance data. *Methods Enzymol.* **563**, 531–567 (2015).

a

```
                                  22              440      469       493
                                  |               |        |         |
xGluN1   1 MGTMRLFLLAVLFLFSFARAGCDPK...GYQMST...NGPNETI...REMNFTY...NSNKKEWNG...
ΔATD     1 MGTMRLFLLAVLFLFSFARAGA---...---MST...NGPDETI...REMDFTY...NSNAAAWNG...

               588    595        610       617        656          769
               |      |          |         |          |            |
xGluN1   ...GRFKVNSEEEEEDA...FSWGVLLNSGIGEG...LVLDRPE...WKQNVSL...TFENMAGVFMLVA
ΔATD     ...GRF--------EDA...FSWRVLLNSGIGEG...LVLRRPE...WKQEVSL...TFENMAGVFMLVA

                              834
                              |
xGluN1   GGIVAGIFLIFIEIA...
ΔATD     GGIVAGIFLIFIEIAYKSRAEAKRMKG
                            A2 tail
```

```
                                           439        490       584        593
                                           |          |         |          |
xGluN2B                      ...DEH...QIRIENR...INGTWNG...GYNRCLADGREPGGPS...
ΔATD     MLGPCMLLLLLLLGLRLQLSLGDEH...QIRPENR...INGVWNG...GYN----------GPS...
         Alkaline phosphatase signal peptide

                  615        654 655        720         772
                  |          |   |          |           |
xGluN2B  ...NNSVPVQ...MIQEEYVD...SLKTGKL...WKRYVDL...HLF...
ΔATD     ...NNSLPVQ...MIQRRYVD...SLKSGKL...WKRQVDL...HLFYKSRAEAKRMKG
                                                           A2 tail
```

b

c

Extended Data Fig. 1 | The ΔATD NMDA receptor construct and structure. a, Selected amino acid sequences of constructs used in these studies are compared to the wild-type sequence to highlight mutations in both subunits. Locations of mutated sites and deletions are highlighted in yellow. Insertions are shown in blue or red. The 'A2 tail' is derived from residues 837–847 of GluA2 AMPA receptor C terminus (NP_058957).

b, Cartoon representation shows the GluN1 and GluN2B subunit constructs and modifications of the ΔATD receptor. The locations of point mutations are highlighted with blue circles and the deletions are shown as yellow wedges. **c**, Superposition of the two ΔATD NMDA receptors in the crystallographic asymmetric unit, aligned by the TMD. Black arrows show the shift between receptor 1 (light blue) and receptor 2 (magenta).
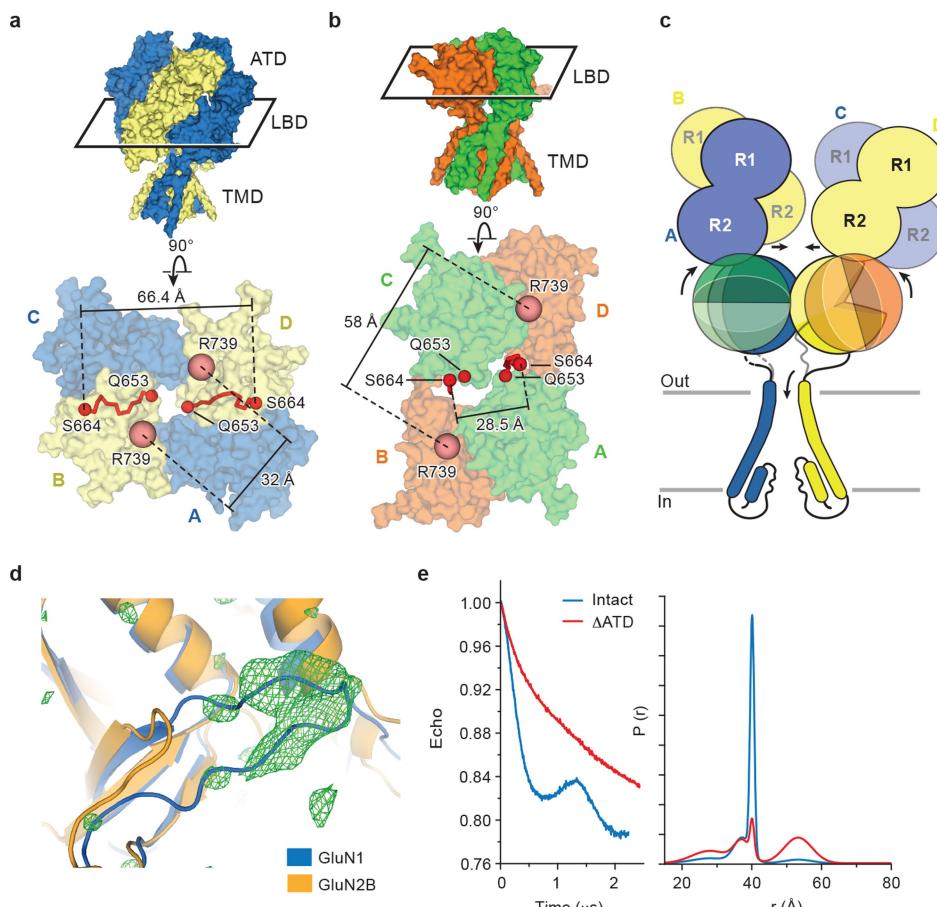
Extended Data Fig. 2 | LBD dimer rearrangement and dynamics.
a, **b**, Top-down view from the extracellular side of the membrane, showing the LBD layer of the intact Δ2 NMDA receptor (**a**) with GluN1 in blue and GluN2B in yellow, and of the ΔATD NMDA receptor (**b**) with GluN1 in green and GluN2B in orange. The M3–LBD linkers of GluN2B (red ribbon, Q653–S664) adopt distinct conformations in the two receptors. Shown are distances between GluN2B R739 residues (β-carbon atoms, salmon spheres), the residue selected for the DEER experiments (in Å) in both the intact and ΔATD receptors. **c**, Cartoon emphasizing how the ATDs participate in defining the conformation of the LBD layer and how this, in turn, keeps the GluN2B M3–D2 linker in a conformation capable of opening the channel gate. **d**, The $F_o - F_c$ density (3σ, green mesh) fits loop 1 of the GluN1 subunit (blue cartoon) but not that of the GluN2B subunit (orange cartoon). **e**, DEER data of MTSSL-labelled

GluN2B(R739C) ΔATD (red) (sample size, $n=2$) and intact NMDA receptor (blue) (sample size, $n=1$). Left, peak-normalized echo decay and fits; right, probability distributions of DEER distances. The probability distributions of the DEER distances show two major peaks, one centred at 35–40 Å and a second broad peak at around 55 Å. The amplitudes of the two peaks in the ΔATD receptor are comparable, with the 55 Å peak corresponding to the 'rearranged' LBD layer as seen in the ΔATD crystal structure, whereas the shorter distance (approximately 35–40 Å) indicates the canonical LBD arrangement, like that observed in the intact receptor structure. The intact receptor, by contrast, shows one major narrow peak at around 40 Å, which corresponds with the predicted distance based on the intact receptor crystal structure, whereas the small peak centred around 55 Å suggests that the intact receptor may harbour a minor population with a ΔATD-like LBD arrangement.
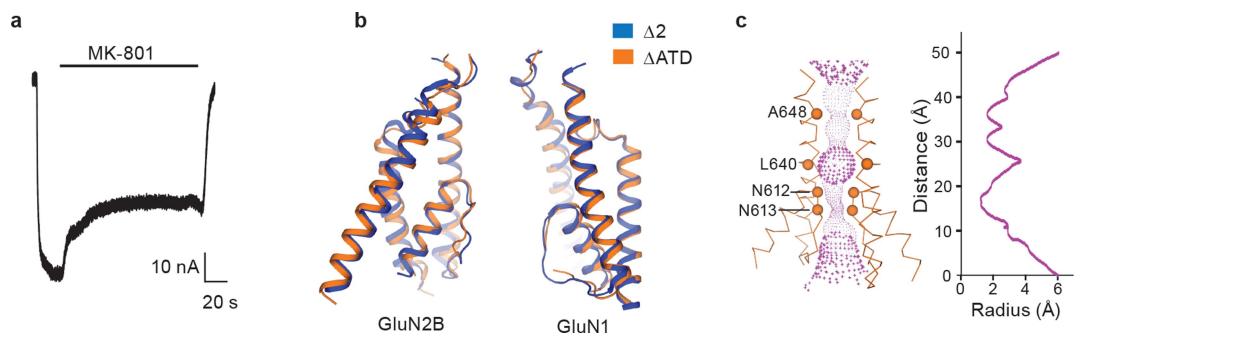
Extended Data Fig. 3 | The ΔATD NMDA receptor channel. a, Inhibition of agonist-induced (300 μM glutamate and 300 μM glycine) current by 1 μM MK-801 for the ΔATD NMDA receptor by two-electrode voltage clamp. The inhibition ratio is 0.37 ± 0.06 (mean ± s.d., $n = 5$). The holding potential is −60 mV. **b**, Superposition of the Δ2 and ΔATD receptor TMDs shows that they adopt similar conformations. **c**, Side view of the ion pore with GluN2B subunits (orange ribbons) showing the van der Waals radius along the pore (magenta dots). The α-carbons of selected residues facing the pore are shown as spheres. The radius is plotted against the distance along the pore axis.

a

b
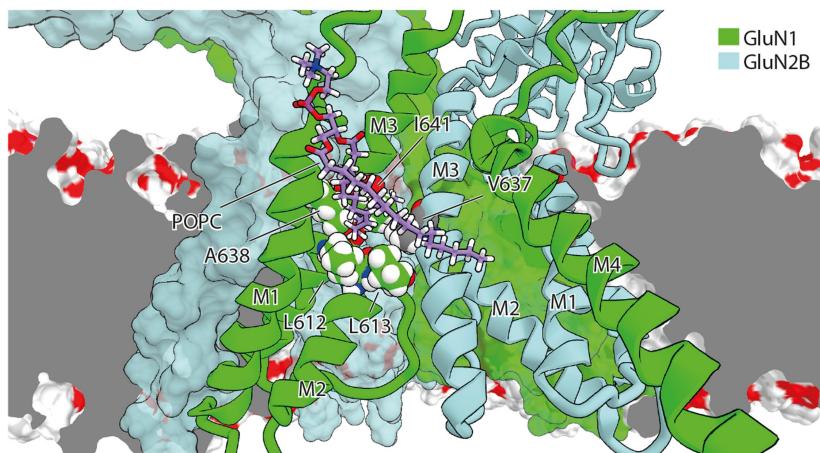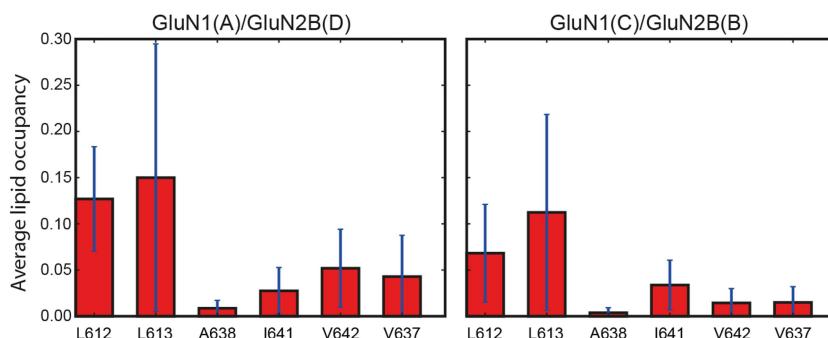
Extended Data Fig. 4 | Lipid accessibility of the TMD tunnel.
a, Simulation snapshot (simulation 2) of a lipid molecule with one of its tails trapped between the M2 and M3 helices of the GluN1 subunit (chain A, green ribbons) and the M3 helix of the adjacent GluN2B subunit (light blue ribbons) viewed from within the membrane and towards the pore. Residues L612, L613, A638, I641 and V642 of GluN1 (chain A) and V637 of GluN2B (chain D) of the tunnel walls are shown as spheres with the carbon atoms in green and grey, respectively. GluN1 (chain C) and GluN2B (chain B) subunits are shown as green and light blue solid surfaces, respectively. The dark grey plane represents a cut across the

lipid membrane, the remainder of which is shown as a red and white surface. **b**, Mean lipid occupancy (number of lipid atoms) within 3.5 Å of the tunnel walls, defined by residues L612, L613, A638, I641 and V642 of GluN1 (chains A, C) and residue V637 of the GluN2B (chains B, D) subunit. The occupancy was calculated across the closed-pore and pore-opening simulations (simulations 2 and 3) and all permeation simulations (simulations 4–17). N, all individual simulations within a given panel; n, all individual data points aggregated across all simulations. $N = 16$ (simulations 2–17); $n \gg 10$. Error bars, s.d. of the mean calculated from all individual data points aggregated across all simulations.
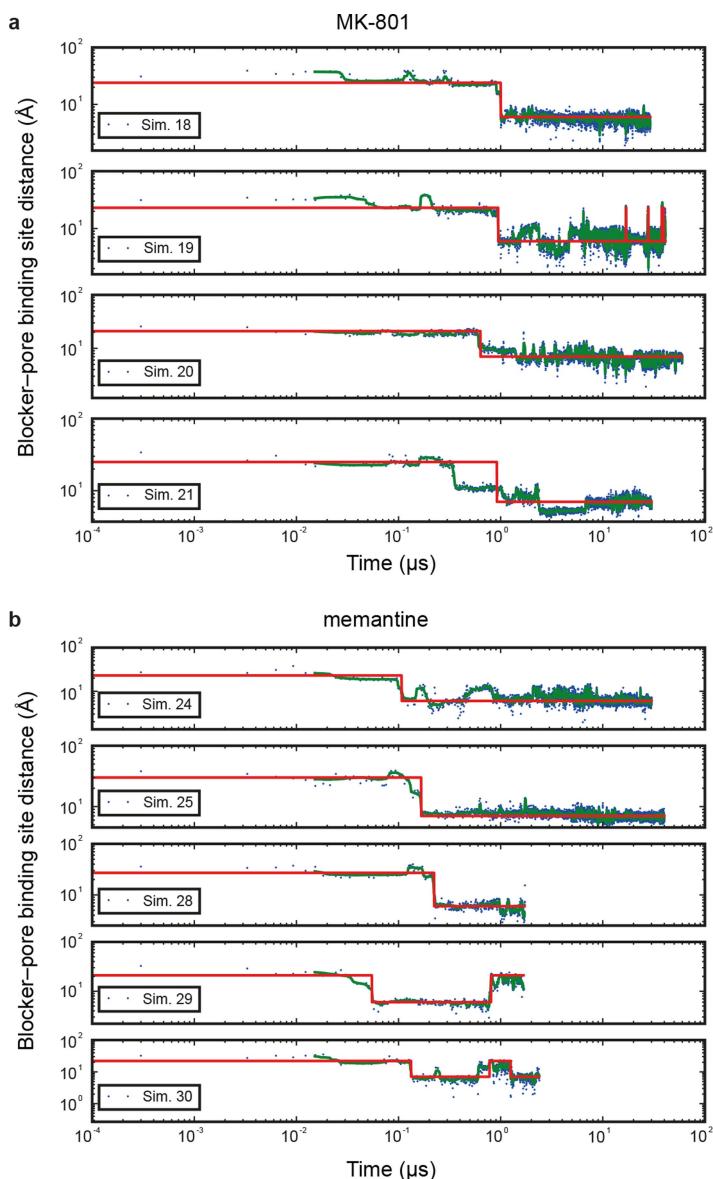
Extended Data Fig. 5 | Binding times of MK-801 and memantine.
a, Binding time of MK-801 in simulations 18–21. Green lines are 30-ns running medians, and red lines indicate bound and unbound states. Binding was defined as the ligand heavy-atom centre of mass being within 10 Å of the centre of mass of the α-atoms of the N612 residues of the two GluN2B subunits. The mean binding time of simulations 20 and 21 at 0 mV was 0.78 ± 0.10 μs; application of voltage in simulations 18 and 19 (-593.9 ± 3.8 and -197.9 ± 1.2 mV) did not significantly decrease the binding time. **b**, Binding time of memantine in simulations 24, 25, and 28–30. Simulations 26 and 27 were initiated with memantine already bound, and the binding curves from these simulations were thus omitted in the determination of the on-rates for this pore blocker. Green lines are 30-ns running medians, and red lines indicate bound and unbound states. The mean binding time in simulations 24 and 25 (at 0 mV) was 0.14 ± 0.02 μs; application of voltage in simulations 28–30 (-592.6 ± 0.3, -592.7 ± 0.3 and -196.9 ± 0.1 mV) did not significantly decrease the binding time. In simulations 29 and 30, 'unbound' states following binding are artefacts owing to the voltage driving memantine through the selectivity filter. $N = 1$ in each panel.
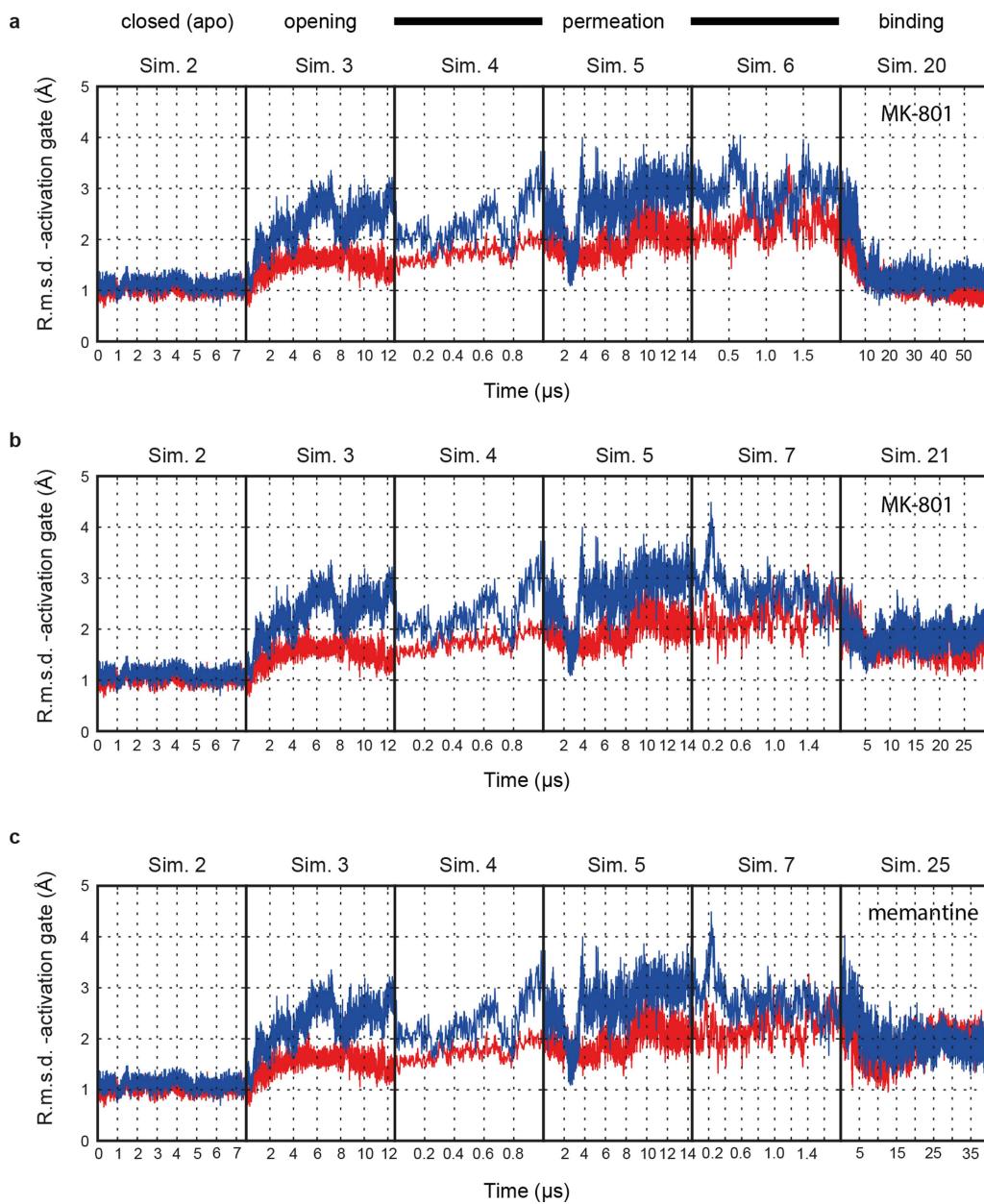
Extended Data Fig. 6 | Blocker-induced channel closure. a–c, The resemblance between the closed, deactivated receptor (far left) and the closed, pore-blocked receptor (far right) is shown. r.m.s.d. (red, GluN1; blue, GluN2B) of the M3-bundle-crossing region (the activation gate) relative to the closed-state Δ2 crystal structure obtained from simulations of the closed pore (simulation 2), pore opening (simulation 3), permeation (simulation 4 at −396.6 ± 2.7 mV, simulation 5 at −593.8 ± 3.8 mV, and simulation 6 (**a**) at −396.1 ± 2.7 mV or simulation 7 (**b** and **c**) at −415.1 ± 6.4 mV), two MK-801-binding simulations (simulations 20 (**a**) and 21 (**b**)), and one memantine-binding simulation (simulation 25 (**c**)). *N* = 1 in each panel.

a

b
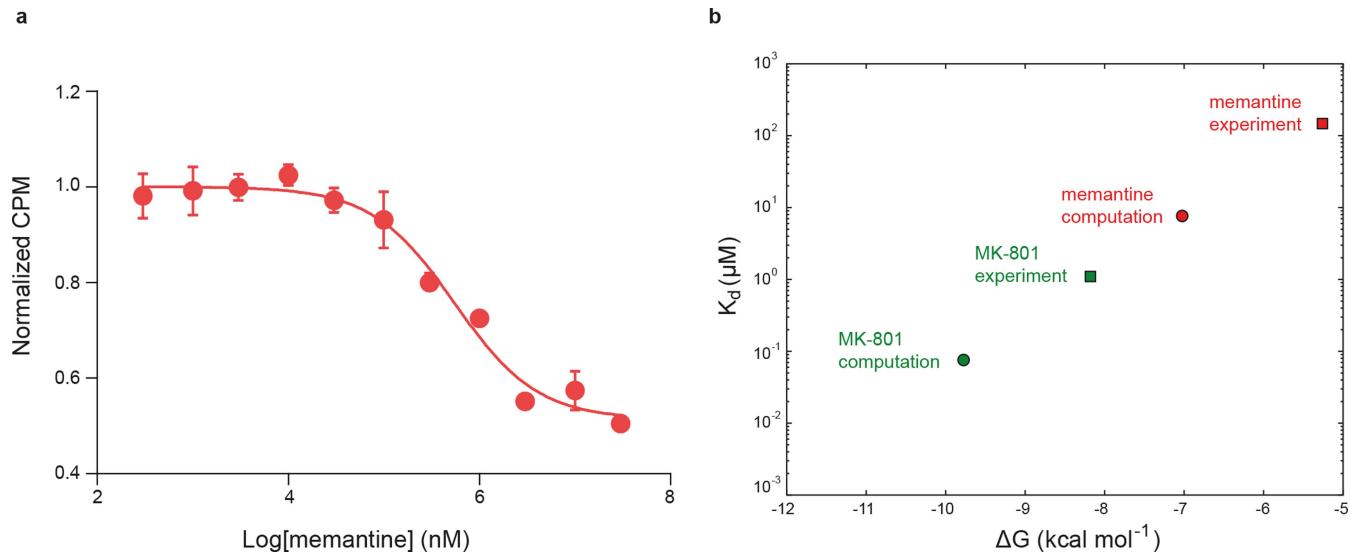
Extended Data Fig. 7 | Free-energy estimates of MK-801 and memantine binding. a, Competition binding of memantine to the $\Delta 2$ receptor in the presence of $3\,\mu M$ [^{3}H]MK-801, measured by the scintillation proximity assay. The plot shows data from a representative experiment with error bars representing s.e.m. from triplicate measurements. **b**, K_d (circles) derived from free-energy estimates of binding of memantine and MK-801 to the open, intact, activated $\Delta 2$ receptor in which the pore has collapsed onto the ligand. The absolute experimental affinities of MK-801 (green) and memantine (red) for the $\Delta 2$ receptors are shown as squares. The free energies were calculated for four independent, ligand-bound configurations, all taken from binding simulations at zero transmembrane voltage (MK-801: simulations 20 and 21; memantine: simulations 24, 25 and 27). Each calculation consisted of $0.5\,\mu s$ of simulation to solvate the ligand in water, followed by $3.0\,\mu s$ of simulation of the protein–ligand complex. The average K_d values for MK-801 and memantine of around 0.08 and around $7.64\,\mu M$ for the $\Delta 2$ receptor show a 100-fold difference in the affinities of these two ligands; a similar relative affinity of these two ligands has been found experimentally using the $\Delta 2$ receptor with a K_d value of around $1.1\,\mu M$ for MK-801 and a K_i value of around $147.4\,\mu M$ for memantine. The calculated binding free energies $-9.78 \pm 1.61\,\mathrm{kcal\,mol}^{-1}$ (MK-801) and $-7.02 \pm 1.24\,\mathrm{kcal\,mol}^{-1}$ (memantine), and thus the free-energy-derived dissociation constants, are subject to large errors, estimated as s.e.m., owing to lack of convergence from including long-range effects from lipid molecules surrounding the pore. We note that the contribution of pore-cavity collapse upon ligand binding to the binding free energy is not included in the free-energy calculations, which were performed with the pore-collapsed, intact, agonist-bound receptor. Also not included is the contribution of $-\ln_2 k_B T$ (where k_B is the Boltzmann constant and T is absolute temperature) arising from the two poses available to MK-801.
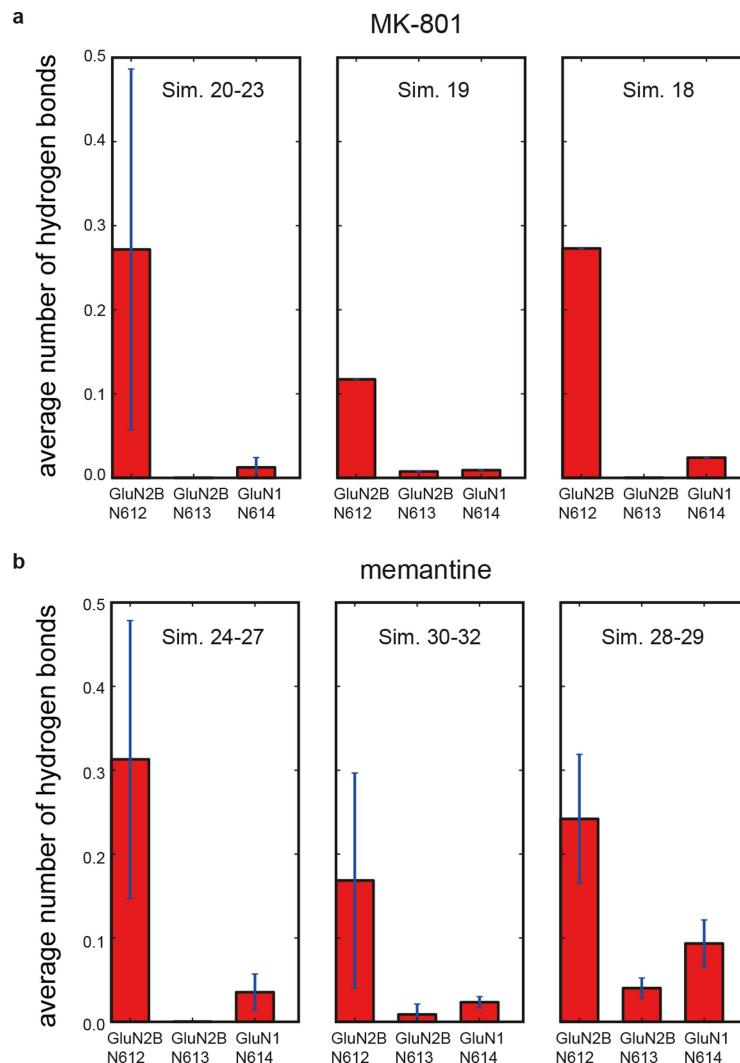
Extended Data Fig. 8 | Hydrogen-bonding propensity between MK-801 and memantine, and the selectivity filter asparagine residues.
a, **b**, The two N-site asparagine residues N614 (GluN1) and N612 (GluN2B) of the pore-loop tips and the N + 1 asparagine residue, N613, of the GluN2B subunit, which is believed to be involved in the voltage-dependence of memantine binding. For MK-801 (**a**): left, data obtained at zero transmembrane voltage (simulations 20–23, $N = 4$); middle, -197.9 ± 1.2 mV (simulation 19, $N = 1$); right, -593.9 ± 3.8 mV (simulation 18, $N = 1$); $n \gg 10$. For memantine (**b**): left, data obtained at zero transmembrane voltage (simulations 24–27, $N = 4$); middle,

-196.9 ± 0.1, -205.1 ± 0.1, and -204.1 ± 0.1 mV (simulations 30–32, $N = 3$); right, -592.6 ± 0.3 and -592.7 ± 0.3 mV (simulations 28 and 29, $N = 2$); $n \gg 10$. Hydrogen-bonding propensity is relatively low for N614 of GluN1, except at high voltage, suggesting that N612 and N613 of GluN2B are more important for pore-blocker binding and its voltage dependence, respectively; at non-zero transmembrane voltage, hydrogen-bonding propensity increases at the N + 1 site asparagine N613 of GluN2B. Error bars, s.d. of the mean calculated from all individual data points aggregated across all simulations.

a MK-801

Sim. 21 Sim. 22 Sim. 23

μ=6.70, σ=2.15   μ=5.80, σ=1.18   μ=6.75, σ=2.12

R.m.s.d.

b memantine

Sim. 25 Sim. 26 Sim. 27

μ=4.15, σ=1.00   μ=4.34, σ=1.01   μ=4.51, σ=0.93

R.m.s.d.

c

M3(D)
M3(B)
M3(A)   M3(C)
N612 (B)
M2(A)
N614 (A,C)   M2(C)
M2(B)

d

M3(C)
M3(D)
M3(B)
M1(D)   M1(B)
N612(D)   M2(C)
M2(D)   N614(C)
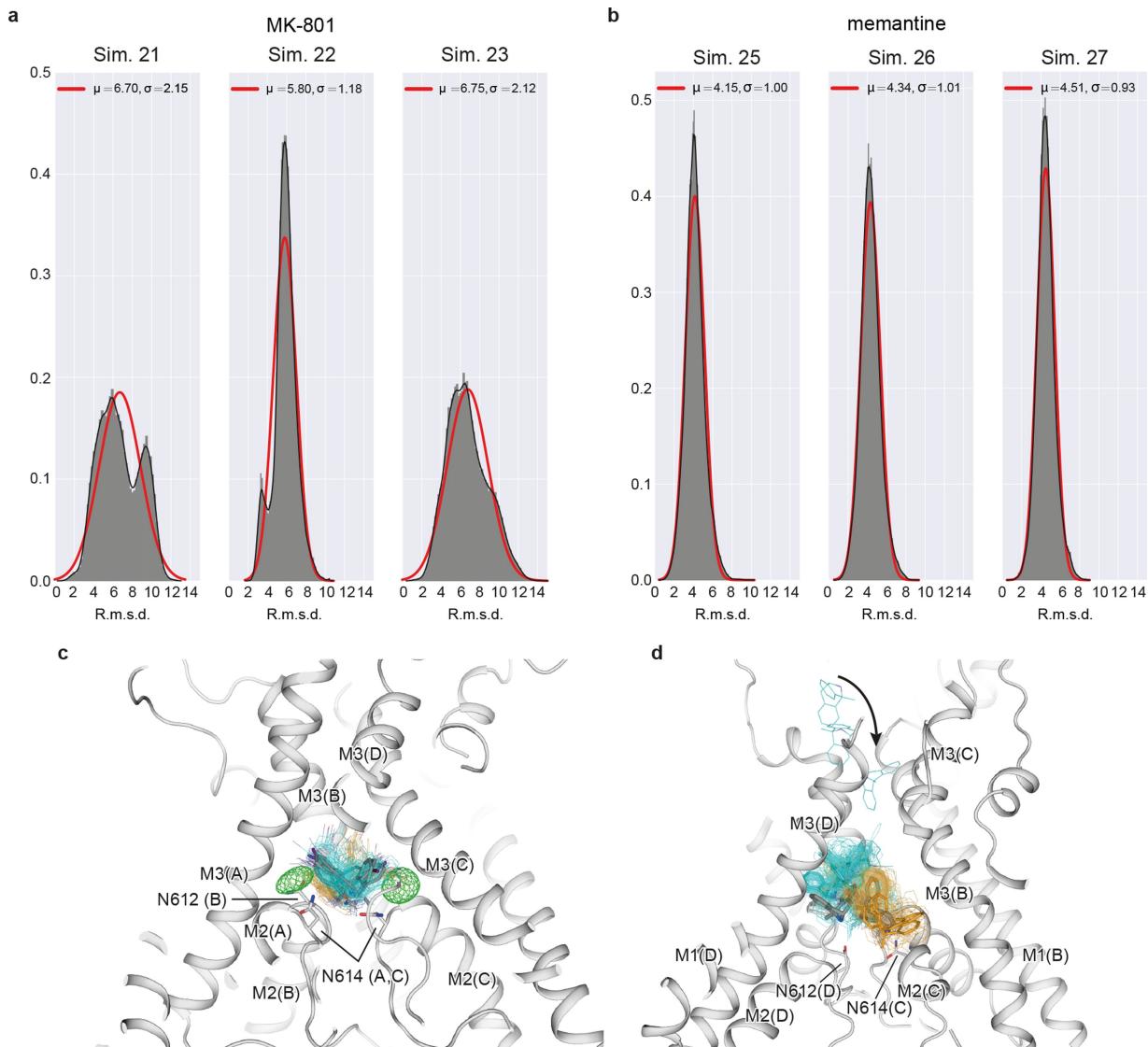
Extended Data Fig. 9 | Binding mode distributions of MK-801 and memantine. a, MK-801 r.m.s.d. distributions (in Å; heavy atoms only) obtained from MK-801-binding simulations 21–23, with respect to all MK-801 poses obtained in binding simulation 20. Mean (μ) and standard deviation (σ) are indicated in Å; solid red lines are best fits to a normal distribution, but the distributions for simulations 21 and 23 show clear evidence of two r.m.s.d. populations, consistent with the observation that MK-801 can block the pore in two symmetry-related poses. The degree of asymmetry of the distributions observed for simulations 21 and 23 indicates non-equal occupancy of the two poses, a result of incomplete sampling. We note that in simulation 22, one of the two poses almost completely predominates. $N = 1$ for each panel; $n \gg 10$. **b**, Memantine r.m.s.d. distributions (in Å; heavy atoms only), obtained from binding simulations 25–27, with respect to all poses obtained in memantine-binding simulation 24. Mean (μ) and standard deviation (σ) are indicated in Å; solid lines are best fits to a normal distribution. The relatively narrow and unimodal distributions reflect that memantine appears to predominantly block the pore in a single pose. The heavy-atom average r.m.s.d. of the main poses of memantine was 3.7 ± 0.2 Å, less than that observed for MK-801. $N = 1$ for each panel; $n \gg 10$. **c**, MK-801-I poses obtained in simulations with and without selectivity filter backbone torsional corrections. Grey, the two predominant poses

observed with corrections (simulation 1); cyan and orange, predominant poses identified from the initial portion (1–3 μs), before the filter deteriorated too extensively, of two different simulations without torsional corrections (pose 1, simulation 42; pose 2, simulation 47). One hundred individual poses from the initial portion (1–3 μs, uniformly separated by 0.02 μs) are shown as cyan and orange lines with the iodine atoms shown as spheres. Both poses of MK-801-I observed in our simulations with torsional backbone corrections (simulation 1) were therefore also observed, with comparable stability, in these additional simulations without these corrections. **d**, MK-801 poses obtained in free-binding simulations with and without filter backbone torsional corrections. Grey, the two distinct poses observed in a free binding simulation with backbone corrections (simulation 20); cyan and orange, poses in the three independent simulations without corrections in which MK-801 bound stably (simulation 50: pose identification period, 4–18 μs; simulation 51: pose identification period, 9–12 μs; simulation 53: pose identification period, 3–4 μs). MK-801 bound stably to the receptor in three (simulations 50, 51 and 53) out of five simulations performed without corrections—again in two distinct poses, as observed in our simulations with torsional corrections—and some closure of the activation gate (that is, the bundle-crossing region) was also observed in these three simulations before the filter deteriorated.

Extended Data Table 1 | Crystallographic and structure refinement statistics

	Dataset 1	Dataset 2
Data collection	APS-IDE	APS-IDC
Space group	$P2_1$	$C2$
Cell dimensions a, b, c (Å)	181.6, 108.5, 182.5	207.5, 120.5, 231.3
Cell angles α, β, γ (°)	90.0, 111.4, 90.0	90.0, 102.8, 90.0
Wavelength (Å)	0.98	1.5
Resolution (Å)*	50-3.58 (3.68-3.58)#	50-5.95 (6.31-5.95)
Completeness*	97.5 (97.9)	97.0 (96.3)
Multiplicity*	5.0 (3.9)	3.5 (3.5)
I/σI*	7.64 (1.91)	9.85 (0.94)
R_{meas} (%)*	10.0 (58.8)	8.0 (164.3)
$CC_{1/2}$ (%)*	99.6 (14.6)	99.8 (50.6)
Anisotropy (Å: a*/b*/c*)#	3.9 / 3.7 / 3.4	
Refinement		
Resolution (Å)	50-3.6 (3.65-3.6)	
No. of reflections	69975	
R_{work}/R_{free} (%)	29.2(35.1)/31.8 (40.1)	
No. of atoms total	20205	
Ligand	222	
Average B-factor (Å^2)		
Protein	124	
Ligand	119	
R.m.s. deviations		
Bond lengths (Å)	0.006	
Bond angles (°)	0.770	
Ramachandran plot		
Favored (%)	93.9	
Allowed (%)	6.1	
Disallowed (%)	0	
Rotamer outliers (%)	4.4	

*Highest resolution shell in parentheses.
#Estimates of anisotropy calculated using the diffraction anisotropy server (http://services.mbi.ucla.edu/anisoscale/).
‡R_{sym} is reported.
5% of reflections were used for calculation of R_{free}.

LETTER

Structural basis of ligand binding modes at the neuropeptide Y Y1 receptor

Zhenlin Yang[1,2,15], Shuo Han[1,3,15], Max Keller[4,15]*, Anette Kaiser[5,15], Brian J. Bender[6,15], Mathias Bosse[7], Kerstin Burkert[5], Lisa M. Kögler[5], David Wifling[4], Guenther Bernhardt[4], Nicole Plank[4], Timo Littmann[4], Peter Schmidt[7], Cuiying Yi[1], Beibei Li[1,8], Sheng Ye[3], Rongguang Zhang[3,9], Bo Xu, Dan Larhammar[10], Raymond C. Stevens[11,12], Daniel Huster[7], Jens Meiler[6,13], Qiang Zhao[1,2,8,14], Annette G. Beck-Sickinger[5]*, Armin Buschauer[4,16] & Beili Wu[1,8,12,14]*

Neuropeptide Y (NPY) receptors belong to the G-protein-coupled receptor superfamily and have important roles in food intake, anxiety and cancer biology[1,2]. The NPY–Y receptor system has emerged as one of the most complex networks with three peptide ligands (NPY, peptide YY and pancreatic polypeptide) binding to four receptors in most mammals, namely the Y1, Y2, Y4 and Y5 receptors, with different affinity and selectivity[3]. NPY is the most powerful stimulant of food intake and this effect is primarily mediated by the Y1 receptor (Y1R)[4]. A number of peptides and small-molecule compounds have been characterized as Y1R antagonists and have shown clinical potential in the treatment of obesity[4], tumour[1] and bone loss[5]. However, their clinical usage has been hampered by low potency and selectivity, poor brain penetration ability or lack of oral bioavailability[6]. Here we report crystal structures of the human Y1R bound to the two selective antagonists UR-MK299 and BMS-193885 at 2.7 and 3.0 Å resolution, respectively. The structures combined with mutagenesis studies reveal the binding modes of Y1R to several structurally diverse antagonists and the determinants of ligand selectivity. The Y1R structure and molecular docking of the endogenous agonist NPY, together with nuclear magnetic resonance, photo-crosslinking and functional studies, provide insights into the binding behaviour of the agonist and for the first time, to our knowledge, determine the interaction of its N terminus with the receptor. These insights into Y1R can enable structure-based drug discovery that targets NPY receptors.

NPY is a highly abundant neuropeptide in the central nervous system[7]. The first characterized NPY receptor Y1R is widely expressed in a variety of tissues and is involved in the regulation of many physiological functions, some of which are known to be related to obesity[8] and cancer[9]. To better understand the ligand-binding behaviour of NPY receptors and provide a basis for drug discovery, we solved crystal structures of Y1R in complex with two structurally diverse antagonists, UR-MK299, an argininamide with high Y1R selectivity[10], and BMS-193885, which displays anorectic activity in animal models[6] (Fig. 1 and Extended Data Table 1). To facilitate the determination of structure, an engineered Y1R construct was designed (see 'Cloning and protein expression' in Methods).

Within the β-branch of class A G-protein-coupled receptors (GPCRs), to which NPY receptors belong, the structures of four receptors, namely the neurotensin receptor NTS1[11], the OX1 and OX2 orexin receptors[12,13] and the endothelin ETB receptor[14], have been

determined so far. These structures reveal differences of ligand-binding modes between different receptors, suggesting that more structural information is needed to develop any consensus about the ligand recognition mechanisms for this GPCR subfamily. The Y1R structure shares a canonical seven-transmembrane helical bundle (helices I–VII) with the other known GPCR structures (Fig. 1 and Extended Data Fig. 1a, b). The Y1R–UR-MK299 and Y1R–BMS-193885 complexes are structurally similar with a C_α root-mean-square deviation (r.m.s.d.) of 0.75 Å within the helical bundle. Both structures exhibit inactive conformations with helix VI adopting an inward conformation that is similar to other inactive GPCR structures. UR-MK299 binds to Y1R in a cavity within the helical bundle bordered by helices III, IV, V, VI and VII (Fig. 2a, b). The diphenylmethyl moiety of the antagonist interacts with a hydrophobic cluster formed by F282[6.54], F286[6.58] and F302[7.35] (superscripts on residues throughout the text indicate Ballesteros–Weinstein nomenclature[15]) on helices VI and VII of Y1R. The critical role of this hydrophobic patch in recognizing the argininamide-type
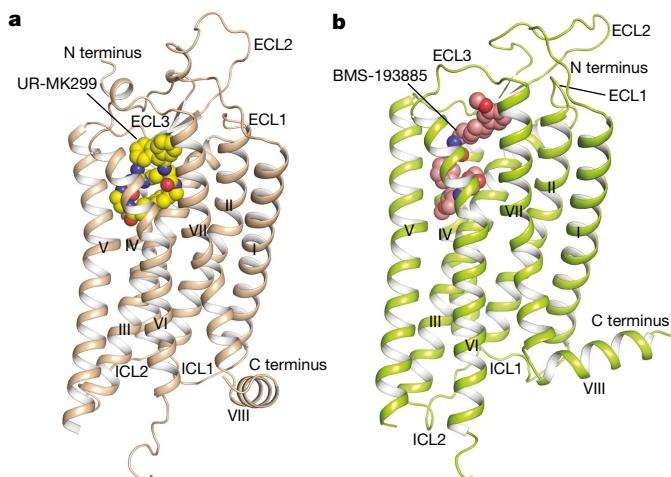
Fig. 1 | Structures of the Y1R–UR-MK299 and Y1R–BMS-193885 complexes. a, Structure of the Y1R–UR-MK299 complex. The receptor is shown in brown cartoon representation. UR-MK299 is shown as spheres with carbons in yellow. b, Structure of the Y1R–BMS-193885 complex. The receptor is shown in green cartoon representation. BMS-193885 is shown as spheres with carbons in pink.

[1]Chinese Academy of Sciences Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [2]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [3]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. [4]Pharmaceutical/Medicinal Chemistry II, Institute of Pharmacy, University of Regensburg, Regensburg, Germany. [5]Institute of Biochemistry, Faculty of Life Sciences, Leipzig University, Leipzig, Germany. [6]Department of Pharmacology, Center for Structural Biology, Vanderbilt University, Nashville, TN, USA. [7]Institute for Medical Physics and Biophysics, Leipzig University, Leipzig, Germany. [8]University of Chinese Academy of Sciences, Beijing, China. [9]National Center for Protein Science Shanghai, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. [10]Department of Neuroscience, Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [11]iHuman Institute, ShanghaiTech University, Shanghai, China. [12]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [13]Departments of Chemistry and Bioinformatics, Center for Structural Biology, Vanderbilt University, Nashville, TN, USA. [14]Chinese Academy of Sciences Center for Excellence in Biomacromolecules, Chinese Academy of Sciences, Beijing, China. [15]These authors contributed equally: Zhenlin Yang, Shuo Han, Max Keller, Anette Kaiser, Brian J. Bender. [16]Deceased: Armin Buschauer. *e-mail: abeck-sickinger@uni-leipzig.de; max.keller@chemie.uni-regensburg.de; beiliwu@simm.ac.cn
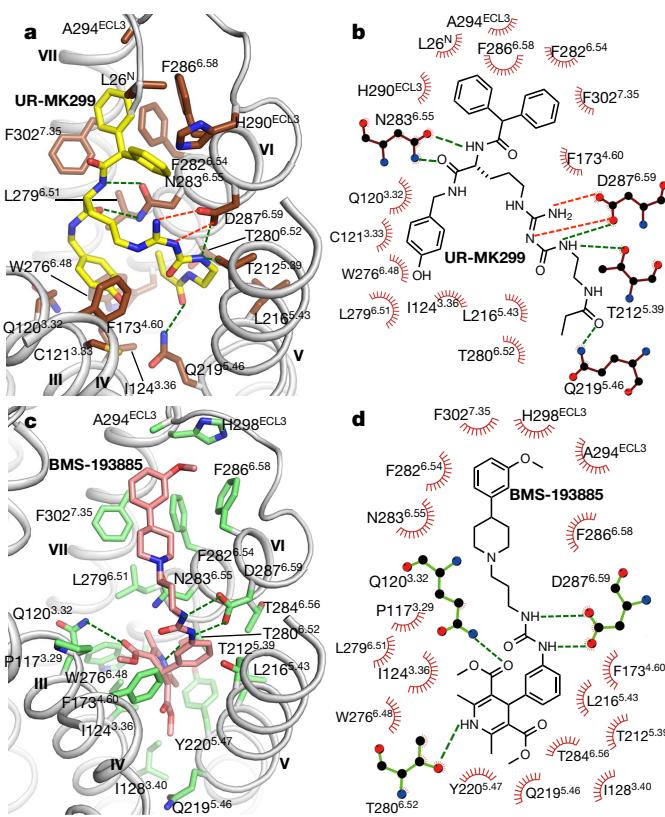
Fig. 2 | Ligand-binding pocket in Y_1R for UR-MK299 and BMS-193885.
a, Binding pocket for UR-MK299. The receptor is shown in grey cartoon representation. UR-MK299 (yellow carbons) and receptor residues (dark brown carbons) involved in ligand binding are shown as sticks. Salt bridge and hydrogen bonds are shown as red and green dashed lines, respectively. **b**, Schematic representation of the interactions between Y_1R and UR-MK299 analysed using the LigPlot$^+$ program[30]. The stick drawing of Y_1R residues is coloured dark brown. **c**, Binding pocket for BMS-193885. BMS-193885 (pink carbons) and receptor residues (green carbons) involved in ligand binding are shown as sticks. **d**, Schematic representation of interactions between Y_1R and BMS-193885 analysed using the LigPlot$^+$ program[30]. The stick drawing of Y_1R residues is coloured green.

Y_1R antagonist was confirmed by the NPY-induced inositol phosphate accumulation of Y_1R when inhibited by UR-MK299 and several related Y_1R antagonists—BIBP3226, BIBO3304, UR-HU404 and UR-MK289 (Extended Data Fig. 1e–i). The F302$^{7.35}$A mutation abolishes the antagonistic activity for all these antagonists, and a two- to fivefold decrease in the antagonistic effect of all tested antagonists was observed for the F286$^{6.58}$A mutation (Fig. 3a–c, Extended Data Fig. 2 and Extended Data Table 2).

The hydroxyphenyl group of UR-MK299 sits in a groove formed by helices III and VI of the receptor, enabling hydrophobic contacts with residues Q120$^{3.32}$, C121$^{3.33}$, I124$^{3.36}$, W276$^{6.48}$ and L279$^{6.51}$. In Y_1R and Y_2R, Q120$^{3.32}$ is suggested to be the interaction partner for the C terminus of NPY and crucial for receptor activation[16]. In the Y_1R–UR-MK299 structure, this residue forms a hydrophobic contact with the phenyl ring of the hydroxyphenyl group in UR-MK299, potentially blocking the binding of Y_1R to NPY. Mutagenesis data show that the Q120$^{3.32}$N mutation does not influence the inhibitory effect of Y_1R antagonists on NPY signalling, but a mutation to histidine increases the antagonistic activity of these ligands (Fig. 3d, e and Extended Data Table 2), suggesting that an additional π-stacking interaction with the antagonist is beneficial at this position. The highly conserved residue W$^{6.48}$ represents the 'toggle switch' and was suggested to trigger receptor activation through a conformational change in various GPCRs[17]. In the Y_1R–UR-MK299 structure, the residue W276$^{6.48}$ is in a conformation that is similar to those observed in other inactive class A GPCR

structures and is distinct from their active-state conformations[18,19]. The hydroxyphenyl group of UR-MK299 forms a hydrophobic contact with W276$^{6.48}$, potentially preventing its activation-related motion thus stabilizing the receptor in an inactive conformation. Compared to the wild-type receptor, the Y_1R mutant W276$^{6.48}$A displayed an over 2,000-fold decrease in its binding affinity to [^{3}H]UR-MK299 (Extended Data Table 3) and reduced the antagonistic activity of the argininamide-type Y_1R antagonists by four- to sevenfold (Fig. 3f and Extended Data Table 2), supporting its important role in antagonist recognition.

Residues N283$^{6.55}$ and D287$^{6.59}$ were suggested as the most important amino acids for Y_1R ligand recognition[20]. In the Y_1R–UR-MK299 structure, N283$^{6.55}$ is engaged in two hydrogen bonds with the α-nitrogen and the carboxylic oxygen next to the hydroxybenzylamine moiety of UR-MK299. D287$^{6.59}$ builds a salt bridge with the protonated guanidinyl moiety and a hydrogen bond with the carbamoyl group, in agreement with a decrease in antagonist affinity when the carbamoyl group was replaced by an alkoxycarbonyl, acyl or alkyl group[21]. The mutants N283$^{6.55}$A and D287$^{6.59}$N displayed a notable loss of NPY-induced receptor function, a complete abolishment of antagonistic activity for the small-molecule antagonists (Fig. 3g, h and Extended Data Table 2), and an over 2,000-fold decrease in the binding affinity of Y_1R to [^{3}H]UR-MK299 (Extended Data Table 3). These data indicate the critical roles of these two residues in agonist and antagonist binding. Additionally, the carbamoyl substituent at the guanidine group binds deep in a sub-pocket shaped by helices V and VI, characterized by hydrophobic contacts with L216$^{5.43}$, T280$^{6.52}$ and N283$^{6.55}$, and a hydrogen bond between the oxygen of the propionyl group and Q219$^{5.46}$. The latter was reflected by a 30-fold decrease in the binding affinity of [^{3}H] UR-MK299 to the Y_1R mutants Q219$^{5.46}$A and Q219$^{5.46}$V (Extended Data Table 3). Extra empty space at the bottom of the sub-pocket is observed in the Y_1R–UR-MK299 structure, suggesting that a larger substituent may be allowed (Extended Data Fig. 1c). This is supported by studies showing that other carbamoylated argininamide-type Y_1R antagonists containing longer carbamoyl chains, such as UR-MK136 (Extended Data Fig. 1j), bind to the receptor with a relatively high affinity[10].

UR-MK299 was reported to exhibit high Y_1R selectivity (Y_2R, inhibition constant (K_i) > 3,000 nM; Y_4R and Y_5R, K_i > 10,000 nM) and specificity compared to two related neuropeptide FF (NPFF) receptors (NPFF$_1R$, K_i = 1,000 nM; NPFF$_2R$, K_i > 3,000 nM)[10]. Sequence alignment reveals that most of the key residues involved in UR-MK299 binding are conserved between Y_1R, the other NPY receptors and the NPFFRs, except for F$^{4.60}$, Q$^{5.46}$, N$^{6.55}$ and F$^{6.58}$ (Extended Data Fig. 3), indicating the importance of these four residues in terms of the selectivity and specificity of UR-MK299. Y_2R is the only NPY receptor with L$^{5.46}$ instead of Q$^{5.46}$, preventing key polar contacts. In Y_4R, E$^{6.58}$ disturbs the F$^{6.54}$-F$^{6.58}$-F$^{7.35}$ hydrophobic patch and probably mediates selectivity, supported by the F$^{6.58}$E mutation in Y_1R reducing binding affinity for BIBP3226[22], which contains the same diphenylmethyl group as UR-MK299. Similarly, hydrophilic residues at key positions impede high-affinity binding at Y_5R (T$^{6.58}$) and NPFF$_2R$ (S$^{6.58}$), while the hydrophobic pocket is preserved in NPFF$_1R$ (L$^{6.54}$-I$^{6.58}$-F$^{7.35}$), although with less bulk, leading to a moderate affinity of BIBP3226 (K_i = 18 nM)[10].

The ligands BMS-193885 and UR-MK299 occupy a similar binding pocket within the helical bundle of Y_1R (Fig. 2c, d and Extended Data Fig. 1k). The dihydropyridine group of BMS-193885 fits in a sub-pocket formed by helices III, V and VI, which aligns with previous structure–activity relationship studies showing that larger substituents at position three of the dihydropyridine ring reduced Y_1R binding affinity[23]. Residue T280$^{6.52}$ forms a hydrogen bond with the nitrogen of the dihydropyridine ring as confirmed by our mutagenesis studies, which showed that the T280$^{6.52}$A mutation decreased the binding affinity of BMS-193885 by about 330-fold (Extended Data Table 3), in agreement with a reported lower affinity N-methylated derivative[24]. Additionally, the dihydropyridine ring makes a hydrophobic contact with residue I124$^{3.36}$, which is consistent with a 400-fold decrease in
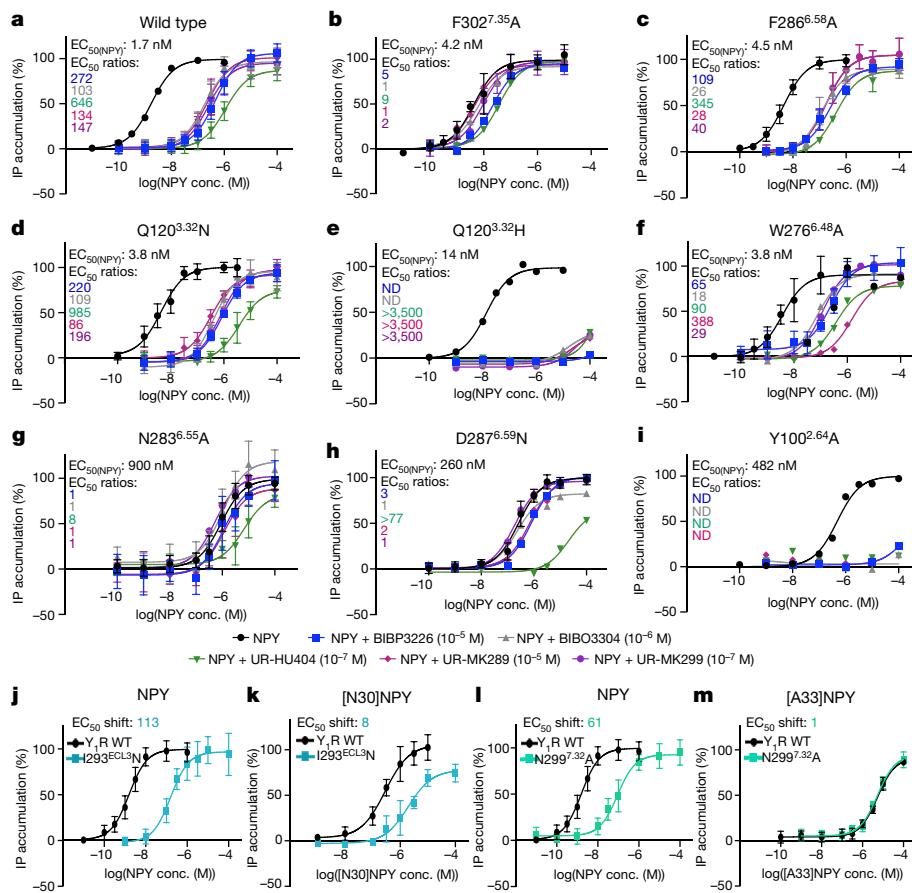
Fig. 3 | Inositol phosphate accumulation assays. a–i, NPY-induced inositol phosphate (IP) accumulation of wild-type (WT) and mutant Y_1 receptors in the absence of antagonists or in the presence of BIBP3226 (10^{-5} M), BIBO3304 (10^{-6} M), UR-HU404 (10^{-7} M), UR-MK289 (10^{-5} M) or UR-MK299 (10^{-7} M). EC_{50} values of NPY (black) and EC_{50} ratios ($EC_{50(NPY + antagonist)}/EC_{50(NPY)}$) for antagonists (coloured) are given in the top left corner for each plot. A reduced EC_{50} ratio of the mutant compared to the wild-type receptor was interpreted as important for the respective antagonist. ND, not determined; NPY conc., concentration of NPY. **j–m**, Complementary mutagenesis assays of [N30]NPY with I293^{ECL3}N

(**j, k**) and [A33]NPY with N299$^{7.32}$A (**l, m**). EC_{50} shifts ($EC_{50(mutant)}/EC_{50(wild type)}$) are given in the top left corner for each plot. A reduced EC_{50} shift of the NPY analogue/Y_1R mutant compared to the NPY/Y_1R mutant was interpreted as no further loss of function and a direct interaction between both positions. At least two independent experiments were performed in technical duplicates. Where more than two experiments were performed (**a–d, f–h, j–m**), data are shown as mean ± s.e.m. Where two experiments were performed (**e, i**), data from a representative experiment are shown. See Extended Data Table 2 for detailed statistical evaluation.

the affinity of the mutant I124$^{3.36}$A (Extended Data Table 3). It was also reported that methylation of either nitrogen of the urea group of BMS-193885 decreased the binding ability of the methylated derivatives to Y_1R[24], suggesting that these hydrogen bond donors are critical for Y_1R recognition. Indeed, in the BMS-193885-bound Y_1R structure, the urea group forms hydrogen bond interactions with D287$^{6.59}$. Similar to the diphenylmethyl group of UR-MK299, the piperidine and methoxyphenyl rings of BMS-193885 form extensive hydrophobic contacts with the residues F282$^{6.54}$, F286$^{6.58}$ and F302$^{7.35}$. Replacement of the methoxyphenyl substituent by piperidine resulted in lower binding affinity to Y_1R[23], indicating the importance of the methoxyphenyl group in Y_1R binding and reflecting lipophilic demands at this position.

Understanding the binding mode of the endogenous agonist NPY at a molecular level will facilitate the rational development of Y_1R-selective ligands. The C-terminal pentapeptide of NPY was found to be essential for binding to the NPY receptors[25]. Because the hydroxyphenyl and the argininamide group of UR-MK299 mimic R35 and Y36 in the C terminus of NPY (Extended Data Fig. 1l), the crystal structure of Y_1R–UR-MK299 serves as a good template for molecular docking of the agonist. To aid docking, complementary mutagenesis studies were performed to determine corresponding interaction partners between Y_1R and NPY (Extended Data Table 4a). Furthermore, solid-state nuclear magnetic resonance (NMR) chemical shift measurements revealed residue-specific alterations of the secondary structure of NPY

upon binding to Y_1R (Extended Data Fig. 4). Several key Y_1R–NPY contacts identified by the mutagenesis studies were used to guide NPY docking in Rosetta[26] with the final models being filtered against the NMR data to generate a final ensemble that best represents the combined data. The NPY-bound model reveals a relatively flat NPY–Y_1R binding pose with the C-terminal tetrapeptide R33–Y36, identified as either a random coil or a β-strand structure in NMR, penetrating into the binding pocket (Fig. 4a). The unstructured N terminus (Y1–P13) is in close proximity to the second extracellular loop (ECL2), while the α-helix in the middle region of NPY (A14–T32) lies along ECL1 and ECL3 and points away from ECL2.

Inspection of the binding pocket of NPY reveals that the binding pose of residue R35 of NPY is similar to that of the argininamide of UR-MK299 (Fig. 4b). R35 forms a salt bridge with the D287$^{6.59}$ residue of Y_1R and approaches N283$^{6.55}$ (Fig. 4c). The NPY mutant R35A displays a decrease in activity of over 6,000-fold, which represents the highest influence on agonist potency of all tested NPY analogues (Extended Data Table 4a), supporting the importance of the positively charged residue in NPY recognition. Aspartate or glutamate residues are not found at position 6.59 in any peptide GPCRs except for the receptors that bind to Arg–Phe-amide peptides, including NPFF, prolactin-releasing peptide and pyroglutamylated Arg–Phe-amide peptide, which share a common C-terminal Arg–Phe–NH$_2$ motif, supporting the hypothesis that the arginine residue may function in a manner that
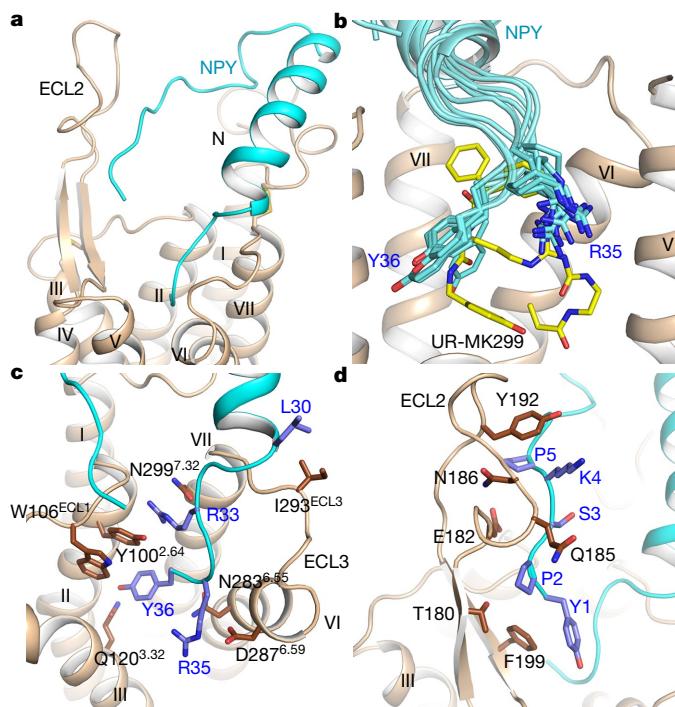
Fig. 4 | Docking poses of NPY. a, Predicted NPY binding pose. The receptor and the lowest energy NPY conformation are shown as cartoons, and are coloured brown and cyan, respectively. **b**, Comparison of UR-MK299 (yellow sticks) binding mode and predicted ensemble binding mode of residues R35 and Y36 of NPY (cyan sticks). **c**, Predicted binding mode between Y_1R and the C terminus of NPY. Key residues involved in the Y_1R–NPY interaction are shown as sticks and coloured dark brown (Y_1R) and blue (NPY). **d**, Predicted binding mode between ECL2 of Y_1R and the N terminus of NPY. Residues in the ECL2 of Y_1R and the N terminus of NPY that may form contacts are shown as dark brown and blue sticks, respectively.

is similar to that of the R35 of NPY by interacting with the conserved $D/E^{6.59}$ of the respective receptors[27]. In contrast to the similarity between the binding modes of the R35 of NPY and the guanidine group of UR-MK299, the C-terminal tyrosinamide of NPY and the hydroxyphenyl group of UR-MK299 show different orientations. The hydroxyphenyl ring is oriented towards helix V ($Q219^{5.46}$) in the UR-MK299-bound Y_1R structure, whereas Y36 of NPY points towards the $Q120^{3.32}$ residue on helix III in the NPY-docked model (Fig. 4c). This may arise from the opposite configuration of the stereo centre in the R35 of NPY and UR-MK299, as well as by only partial mimicking of the $Y36–NH_2$ of NPY by a 4-hydroxybenzyl group in UR-MK299 (Extended Data Fig. 1g, l). In the Y_1R–UR-MK299 structure, residue $Q120^{3.32}$ forms a hydrophobic contact with the hydroxyphenyl group of the antagonist. By contrast, the NPY-bound model shows that the side chain of $Q120^{3.32}$ points almost in the opposite direction and engages in a hydrogen bond with the hydroxy group of $Y36–NH_2$ (Extended Data Fig. 1d), in a similar manner to the previously suggested interaction between the Y_2R residue $Q^{3.32}$ and NPY[28]. In Y_2R, it was also reported that $Q^{3.32}$ may interact with the C-terminal amide of NPY[16]. Inositol phosphate accumulation studies show that the Y_1R mutation $Q120^{3.32}H$ leads to a 26-fold decrease in the potency of NPY, and NPY–tyramide lacking the C-terminal amide displays a 45-fold loss of activity. Complementary mutagenesis analysis revealed an additional reduction of NPY–tyramide potency at the $Q120^{3.32}H$ mutant, and thus rules out a direct contact between the C-terminal amide of NPY and $Q120^{3.32}$ in Y_1R (Extended Data Table 4a). Additionally, Y36 of NPY forms hydrophobic contacts with $Y100^{2.64}$ and $W106^{ECL1}$ in Y_1R (Fig. 4c). Although $Y100^{2.64}$ is not involved in antagonist binding, mutagenesis data suggest a critical role in agonist recognition as the Y_1R mutant $Y100^{2.64}A$ displays a

284-fold decrease in potency for NPY (Fig. 3i and Extended Data Table 2). Furthermore, the model reveals close contacts between L30 of NPY and I293 in ECL3 of Y_1R and between R33 of NPY and the Y_1R residue $N299^{7.32}$ (Fig. 4c), which align with complementary mutagenesis data showing no further loss of function for combining mutant I293N with [N30]NPY and $N299^{7.32}A$ with [A33]NPY (Fig. 3j–m and Extended Data Table 4a).

Previous studies have shown that different NPY receptors behave differently when binding to the N terminus of NPY. Y_2R and Y_5R can bind to N-terminally truncated NPY, whereas Y_1R and Y_4R require the complete N terminus of NPY for full agonist potency[25,27]. However, these data did not allow conclusions about the interaction of the N terminus of NPY with the receptor. To further explore the involvement of the NPY N terminus in recognition between the receptor and ligand, we performed mutagenesis studies, showing that truncation of the first two residues of NPY (NPY(3–36)) reduces peptide potency by more than 50-fold (Extended Data Table 4b). This decrease in potency, however, is not seen when these residues are mutated to alanine ([A1,A2]NPY, fivefold shift in the half-maximal effective concentration (EC_{50})), suggesting important contributions of the peptide backbone in binding to the receptor. Our NPY-bound model suggests that the N-terminal region of NPY makes close contacts with the T180–F199 fragment in ECL2 of Y_1R and is also in proximity to the receptor N terminus (Fig. 4a, d). To experimentally verify interacting sites in the receptor, we performed photo-crosslinking studies between NPY analogues carrying the highly reactive p-benzoyl-phenylalanine [Bpa^1, $K^4[(Ahx)_2$-biotin]]NPY (in which Ahx denotes aminohexanoate) and Y_1R. Crosslinked fragments were assigned to two regions in Y_1R, the N terminus (K21–D32) and the ECL2 (A191–D205) (Extended Data Fig. 5 and Extended Data Table 5). Previous studies demonstrated that deletion of the Y_1R N terminus does not interfere with receptor signalling, but reduces NPY binding by about 95% compared to the full-length receptor[29]. This creates the possibility that the N terminus of Y_1R has a role in recognizing and positioning the peptide ligand, which is in agreement with the photo-crosslinking data. Consistent with the crosslinking hits in receptor ECL2, our mutagenesis data show that the F184A/N and V197N mutations in this region greatly reduce NPY potency (Extended Data Table 4c). Together, these data suggest that the N terminus and ECL2 of Y_1R have important roles in the recognition of the N terminus of NPY and receptor activation. This contrasts with NPY binding at Y_2R, in which ECL2 may interact with the central α-helix of NPY and the peptide N terminus is flexible and not anchored by the receptor[16]. Although this study provides insights into the interactions between Y_1R and NPY, further structural details, such as the structures of Y receptors bound to NPY, are required to fully understand the endogenous agonist-binding modes of the NPY receptor family.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0046-x.

1. Zhang, L., Bijker, M. S. & Herzog, H. The neuropeptide Y system: pathophysiology and therapeutic implications in obesity and cancer. *Pharmacol. Ther.* **131**, 91–113 (2011).
2. Morales-Medina, J. C., Dumont, Y. & Quirion, R. A possible role of neuropeptide Y in depression and stress. *Brain Res.* **1314**, 194–205 (2010).
3. Michel, M. C. et al. XVI. International Union of Pharmacology recommendations for the nomenclature of neuropeptide Y, peptide YY, and pancreatic polypeptide receptors. *Pharmacol. Rev.* **50**, 143–150 (1998).
4. Yulyaningsih, E., Zhang, L., Herzog, H. & Sainsbury, A. NPY receptors as potential targets for anti-obesity drug development. *Br. J. Pharmacol.* **163**, 1170–1202 (2011).
5. Sousa, D. M., Herzog, H. & Lamghari, M. NPY signalling pathway in bone homeostasis: Y1 receptor as a potential drug target. *Curr. Drug Targets* **10**, 9–19 (2009).

6. Antal-Zimanyi, I. et al. Pharmacological characterization and appetite suppressive properties of BMS-193885, a novel and selective neuropeptide Y_1 receptor antagonist. *Eur. J. Pharmacol.* **590**, 224–232 (2008).
7. Leibowitz, S. F., Sladek, C., Spencer, L. & Tempel, D. Neuropeptide Y, epinephrine and norepinephrine in the paraventricular nucleus: stimulation of feeding and the release of corticosterone, vasopressin and glucose. *Brain Res. Bull.* **21**, 905–912 (1988).
8. MacNeil, D. J. NPY Y_1 and Y_5 receptor selective antagonists as anti-obesity drugs. *Curr. Top. Med. Chem.* **7**, 1721–1733 (2007).
9. Reubi, J. C., Gugger, M., Waser, B. & Schaer, J. C. Y. Y_1-mediated effect of neuropeptide Y in cancer: breast carcinomas as targets. *Cancer Res.* **61**, 4636–4641 (2001).
10. Keller, M. et al. N^ω-carbamoylation of the argininamide moiety: an avenue to insurmountable NPY Y_1 receptor antagonists and a radiolabeled selective high-affinity molecular tool ([^{3}H]UR-MK299) with extended residence time. *J. Med. Chem.* **58**, 8834–8849 (2015).
11. White, J. F. et al. Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).
12. Yin, J. et al. Structure and ligand-binding mechanism of the human OX_1 and OX_2 orexin receptors. *Nat. Struct. Mol. Biol.* **23**, 293–299 (2016).
13. Yin, J., Mobarec, J. C., Kolb, P. & Rosenbaum, D. M. Crystal structure of the human OX_2 orexin receptor bound to the insomnia drug suvorexant. *Nature* **519**, 247–250 (2015).
14. Shihoya, W. et al. Activation mechanism of endothelin ET_B receptor by endothelin-1. *Nature* **537**, 363–368 (2016).
15. Ballesteros, J. A. & Weinstein, H. in *Methods in Neurosciences* Vol. 25 (ed. S. Sealfon) 366–428 (Elsevier, Amsterdam, 1995).
16. Kaiser, A. et al. Unwinding of the C-terminal residues of neuropeptide Y is critical for Y_2 receptor binding and activation. *Angew. Chem. Int. Edn Engl.* **54**, 7446–7449 (2015).
17. Venkatakrishnan, A. J. et al. Molecular signatures of G-protein-coupled receptors. *Nature* **494**, 185–194 (2013).
18. Rasmussen, S. G. et al. Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–555 (2011).
19. Standfuss, J. et al. The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **471**, 656–660 (2011).
20. Sautel, M. et al. Neuropeptide Y and the nonpeptide antagonist BIBP 3226 share an overlapping binding site at the human Y_1 receptor. *Mol. Pharmacol.* **50**, 285–292 (1996).
21. Keller, M. et al. Guanidine–acylguanidine bioisosteric approach in the design of radioligands: synthesis of a tritium-labeled N^G-propionylargininamide ([^{3}H]-UR-MK114) as a highly potent and selective neuropeptide Y Y_1 receptor antagonist. *J. Med. Chem.* **51**, 8168–8172 (2008).
22. Sjödin, P. et al. Re-evaluation of receptor-ligand interactions of the human neuropeptide Y receptor Y1: a site-directed mutagenesis study. *Biochem. J.* **393**, 161–169 (2006).
23. Poindexter, G. S. et al. Dihydropyridine neuropeptide Y Y_1 receptor antagonists. *Bioorg. Med. Chem. Lett.* **12**, 379–382 (2002).
24. Poindexter, G. S. et al. Dihydropyridine neuropeptide Y Y_1 receptor antagonists 2: bioisosteric urea replacements. *Bioorg. Med. Chem.* **12**, 507–521 (2004).
25. Pedragosa-Badia, X., Stichel, J. & Beck-Sickinger, A. G. Neuropeptide Y receptors: how to get subtype selectivity. *Front. Endocrinol. (Lausanne)* **4**, 5 (2013).
26. Bender, B. J. et al. Protocols for molecular modeling with Rosetta3 and RosettaScripts. *Biochemistry* **55**, 4748–4763 (2016).
27. Merten, N. et al. Receptor subtype-specific docking of Asp$^{6.59}$ with C-terminal arginine residues in Y receptor ligands. *J. Biol. Chem.* **282**, 7543–7551 (2007).
28. Xu, B. et al. Mutagenesis and computational modeling of human G-protein-coupled receptor Y2 for neuropeptide Y and peptide YY. *Biochemistry* **52**, 7987–7998 (2013).
29. Lindner, D., Walther, C., Tennemann, A. & Beck-Sickinger, A. G. Functional role of the extracellular N-terminal domain of neuropeptide Y subfamily receptors in membrane integration and agonist-stimulated internalization. *Cell. Signal.* **21**, 61–68 (2009).
30. Laskowski, R. A. & Swindells, M. B. LigPlot$^+$: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).

Author contributions Z.Y. and S.H. optimized the construct, developed the purification procedure, purified the Y_1R protein for crystallization, performed crystallization trials, solved the structures and wrote the manuscript. M.K., D.W., G.B., N.P. and T.L. synthesized the compounds, designed, performed and analysed the ligand-binding assay. A.K., K.B. and L.M.K. performed peptide synthesis, inositol phosphate accumulation assays, the photo-crosslinking assay and mass spectrometry after crosslinking. B.J.B. helped to refine the Y_1R–UR-MK299 structure and modelled the Y_1R–NPY complex. M.B. and P.S. performed NMR analysis and analysed NMR data. C.Y. expressed the Y_1R proteins. B.L. helped with construct and crystal optimization. S.Y., R.Z., B.X., D.L., R.C.S., D.H., J.M., A.G.B.-S. and A.B. helped with structure analysis, interpretation and edited the manuscript. R.C.S. helped to initiate the project. D.H. oversaw NMR studies. J.M. oversaw molecular docking. Q.Z. collected X-ray diffraction data and solved the structures. A.G.B.-S. oversaw peptide synthesis, inositol phosphate accumulation and photo-crosslinking assays. A.B. oversaw compound synthesis and ligand-binding assays. B.W. and Q.Z. initiated the project, planned and analysed experiments, supervised the research and wrote the manuscript with input from all co-authors.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0046-x.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0046-x.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to M.K., A.G.B. or B.W.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Cloning and protein expression. The DNA sequence of wild-type human Y$_1$R was optimized and synthesized by Genewiz and then cloned into a modified pFastBac1 vector (Invitrogen), which contains an expression cassette with a haemagglutinin signal sequence followed by a Flag tag before the receptor at the N terminus and a PreScission protease site followed by a 10 × His-tag at the C terminus. An engineered construct was generated by inserting a modified T4 lysozyme (T4L)[31] at the third intracellular loop (ICL3) between residues R241 and D250 and introducing a mutation (F129$^{3.41}$W)[32]. Twenty-five amino acids (V359–I384) were truncated at the C terminus to further improve protein yield and stability. Bac-to-Bac Baculovirus Expression System (Invitrogen) was used to generate high-titre ($> 10^8$ viral particles per ml) recombinant baculovirus. *Spodoptera frugiperda* (Sf9) cells (Invitrogen) at a density of 2×10^6 cells per ml were infected by viral stock at an MOI (multiplicity of infection) of 5. As well as the virus, a ligand (UR-MK299 or BMS-193885) was added to the cell culture to a final concentration of 1 μM. Transfected cells were cultured at 27 °C for 48 h and then collected by centrifugation and stored at −80 °C until use.

Purification of Y$_1$R–UR-MK299 and Y$_1$R–BMS-193885 complexes. Frozen insect cells expressing the Y$_1$R–UR-MK299 complex were disrupted with thawing and repeated dounce homogenization in a hypotonic buffer containing 10 mM HEPES, pH 7.5, 10 mM MgCl$_2$, 20 mM KCl and protease inhibitor cocktail (Roche). After centrifugation at 160,000g for 30 min, cell debris was re-suspended in a high-osmotic buffer (10 mM HEPES, pH 7.5, 1 M NaCl, 10 mM MgCl$_2$, 20 mM KCl) and then homogenized extensively. Soluble and membrane associated proteins were removed from the suspension by centrifugation. This procedure was repeated two to three more times and then the hypotonic buffer was used to remove the high concentration of NaCl. Purified membranes were re-suspended in the hypotonic buffer with additional 30% (v/v) glycerol and stored at −80 °C until use.

Purified membranes were thawed on ice in the presence of 100 μM UR-MK299, 2 mg ml^{-1} iodoacetamide (Sigma) and EDTA-free protease inhibitor cocktail (Roche) and incubated at 4 °C for 1 h. Equal volume of solubilization buffer containing 100 mM HEPES, pH 7.5, 1 M NaCl, 1% (w/v) n-dodecyl-β-D-maltopyranoside (DDM, Anatrace), 0.2% (w/v) cholesterol hemisuccinate (CHS, Sigma) was added and incubation was continued for an additional 3 h. The supernatant was isolated by centrifugation at 160,000g for 30 min and incubated with TALON resin (Clontech) supplemented with 10 mM imidazole, pH 7.5 at 4 °C overnight. The resin was first washed with ten column volumes of 25 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, 30 mM imidazole and 50 μM UR-MK299, then with ten column volumes of 50 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, 10 mM MgCl$_2$, 5 mM ATP and 50 μM UR-MK299 and finally with five column volumes of 25 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol and 50 μM UR-MK299. The protein was eluted by five column volumes of 25 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, 300 mM imidazole and 100 μM UR-MK299. A PD MiniTrap G-25 column (GE healthcare) was used to remove imidazole. The C-terminal His-tag and glycosylation was then treated with overnight digestion with His-tagged PreScission protease (custom-made) and His-tagged PNGase F (custom-made). Ni-NTA super flow resin (Qiagen) reverse binding was performed to remove the PreScission protease, PNGase F and the cleaved His-tag. The purified Y$_1$R–UR-MK299 complex was collected and concentrated to 20–30 mg ml^{-1} with a 100 kDa molecular mass cut-off concentrator (Sartorius Stedim Biotech). Receptor purity and monodispersity were estimated by SDS–PAGE and analytical size-exclusion chromatography.

The Y$_1$R–BMS-193885 complex protein was purified following the same procedure as above. The membranes of the Y$_1$R construct were incubated with 50 μM BMS-193885, 2 mg ml^{-1} iodoacetamide (Sigma), and EDTA-free protease inhibitor cocktail (Roche) at 4 °C for 1 h, and then solubilized in final concentration of 50 mM HEPES, pH 7.5, 500 mM NaCl, 0.5% (w/v) DDM, 0.1% (w/v) CHS, 10% glycerol and 25 μM BMS-193885 at 4 °C for 3 h. The solubilized Y$_1$R–BMS-193885 complex bound to the TALON resin was first washed with ten column volumes of 25 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, 30 mM imidazole and 25 μM BMS-193885, and then with ten column volumes of 50 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, 10 mM MgCl$_2$, 5 mM ATP and 25 μM BMS-193885 and finally with five column volumes of 25 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol and 25 μM BMS-193885. The protein was eluted by five column volumes of 25 mM HEPES, pH 7.5, 500 mM NaCl, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 10% (v/v) glycerol, 300 mM imidazole and 50 μM BMS-193885. The eluted sample was concentrated and desalted using the PD MiniTrap G-25 column (GE healthcare). Overnight

digestion by PreScission protease and PNGase F and Ni-NTA reverse binding were then performed to further purify the protein. The complex protein was concentrated to 10–20 mg ml^{-1} and analysed by SDS–PAGE and analytical size-exclusion chromatography.

Lipidic cubic phase crystallization of antagonist-bound Y$_1$R receptors. The Y$_1$R sample in complex with UR-MK299 or BMS-193885 was mixed with molten lipid (monoolein and cholesterol 10:1 by mass) at a weight ratio of 1:1.5 (protein:lipid) using two syringes to create a lipidic cubic phase (LCP). The mixture was dispensed onto glass sandwich plates (Shanghai FAstal BioTech) in 40 nl drop and overlaid with 800 nl precipitant solution using a Gryphon robot (Art-Robbins). Protein reconstitution in LCP and crystallization trials were performed at room temperature (19–22 °C). Plates were placed in an incubator (Rock Imager, Formulatrix) and imaged at 20 °C automatically following a schedule. Crystals of Y$_1$R–UR-MK299 complex appeared after 4 days and grew to full size ($150 \times 50 \times 5\,\mu m^3$) within two weeks in 0.1 M Tris, pH 7.4–8.0, 30–40% (v/v) PEG400, 50–150 mM sodium tartrate and 100 μM UR-MK299. The Y$_1$R–BMS-193885 complex was crystallized in 0.1 M HEPES, pH 7.2–7.6, 20% PEG400 and 50 μM BMS-193885 with the maximum size of $30 \times 10 \times 5\,\mu m^3$. The crystals of Y$_1$R–UR-MK299 and Y$_1$R–BMS-193885 complexes were harvested directly from LCP using 150 μm and 50 μm micro mounts (M2-L19-50/100, MiTeGen), respectively, and flash frozen in liquid nitrogen.

Data collection and structure determination. X-ray diffraction data were collected at the SPring-8 beam line 41XU, Hyogo, Japan, on a Pilatus3 6 M detector (X-ray wavelength 1.0000 Å). Crystals were exposed with a 10 μm × 8 μm mini-beam for 0.2 s and 0.2° oscillation per frame. Data from the 47 best-diffracting crystals of the Y$_1$R–UR-MK299 complex and 33 crystals of the Y$_1$R–BMS-193885 complex were processed by XDS[33]. The structure of the Y$_1$R–UR-MK299 complex was solved by molecular replacement implemented in Phaser[34] using the receptor portion of NTS1 (Protein Data Bank (PDB) accession number 4GRV), converted to polyalanines, and T4L structure (PDB accession number 1C6P) as search models. The correct molecular replacement solution contained one Y$_1$R-T4L molecule in the asymmetric unit. Initial refinement was performed with REFMAC5[35] and BUSTER[36], and then manual examination and rebuilding of the refined coordinates were carried out in COOT[37] using both $|2F_o| - |F_c|$ and $|F_o| - |F_c|$ maps. The structure has been carefully refined and the Ramachandran plot analysis indicates that 100% of the residues are in favourable (95.5%) or allowed (4.5%) regions (no outliers). The final model includes 306 residues (F18–R241 and S256–F337) of the 384 residues of Y$_1$R and residues N2–Y161 of T4L. The Y$_1$R–BMS-193885 complex structure was solved using Y$_1$R in the Y$_1$R–UR-MK299 complex and T4L as search models and refined using the same procedure. The Ramachandran plot analysis indicates that 100% of the residues are in favourable (95.4%) or allowed (4.6%) regions (no outliers). The final model of the Y$_1$R–BMS-193885 complex contains 301 residues (D31–R241 and D250–D339) of Y$_1$R and the 160 residues of T4L. Helix VIII in the Y$_1$R–UR-MK299 structure rotates towards helix VI by about 90° compared to the BMS-193885-bound structure, this is probably caused by crystal packing (Extended Data Fig. 1).

Immunoblotting. The total solubilized protein of the Sf9 membrane preparations (see above) used in the radio ligand binding assay was determined using the Bradford method according to the manufacturer's protocol (BioRad Protein Assay; BioRad). Aliquots of homogenized membrane preparations, corresponding to 100 μg of protein, were centrifuged at 50,000 g at 4 °C for 15 min, and the pellets were re-suspended in 50 mM Tris, pH 7.4, supplemented with 1 mM EDTA and protease inhibitors (SIGMAFAST Protease Inhibitor cocktail tablets, Sigma) at a protein concentration of 1,600 μg ml^{-1}. Membrane homogenates (15 μl) were processed and subjected to immunoblotting as described previously[38] with the following modifications: blotting onto the nitrocellulose membrane was performed at 60 mA for 60 min. Primary antibody anti-Flag M1 from mouse (Sigma, F3040, lot SLBK1592V) was diluted 1:500. The secondary antibody, an anti-mouse IgG horseradish peroxidase (HRP)-conjugated antibody from goat (Sigma, A0168, lot 080M4839) was diluted 1:80,000. The washing steps after incubation with the primary and the secondary antibody were 3×10 min each. Control experiments in the absence of the primary antibody were not performed.

Radioligand-binding assay. All binding experiments with [^{3}H]UR-MK299 (synthesis described elsewhere[10]) were performed on Sf9 membrane preparations in PP 96-well microplates (Greiner Bio One) at 23 ± 1 °C using a sodium-containing, iso-osmotic HEPES buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM KCl, 2.5 mM CaCl$_2$, 1.2 mM KH$_2$PO$_4$, 1.2 mM Mg$_2$SO$_4$ and 25 mM NaHCO$_3$ supplemented with 1% BSA) for competition-binding studies with antagonists, and a sodium-free, hypo-osmotic HEPES buffer (25 mM HEPES, pH 7.4, 2.5 mM CaCl$_2$ and 1 mM MgCl$_2$ supplemented with 1% BSA) for competition binding studies with the agonist NPY (hereafter, both buffers are referred to as 'binding buffer'). Before competition binding experiments, dissociation constant (K_d) values of [^{3}H] UR-MK299 were determined by saturation binding using the respective binding buffer. In the case of saturation-binding experiments, [^{3}H]UR-MK299 was 1:1

diluted with 'cold' UR-MK299 (hereafter, the mixture is referred to as 'radioligand'). On the day of the experiment, Sf9 membranes were thawed, re-suspended using a 1-ml syringe equipped with a needle (20 G) and then centrifuged at 16,000g at 4 °C for 10 min. The supernatant was discarded and the pellets were re-suspended in binding buffer using a 1-ml syringe equipped with a needle (27G3/4). The membrane homogenates were stored on ice until use. The total amount of protein per well was between 0.25 and 8 μg, depending on the receptor expression level.

Saturation binding experiments. For the determination of total binding, wells were pre-filled with binding buffer (160 μl), and then 20 μl of binding buffer, containing the radioligand at a concentration tenfold higher than the final concentration, was added. For the determination of unspecific binding (in the presence of UR-MK299 at a 100-fold excess), wells were pre-loaded with binding buffer (140 μl), binding buffer (20 μl) containing UR-MK299 (tenfold concentrated) and binding buffer (20 μl) containing the radioligand (tenfold concentrated). To all wells, 20 μl of the membrane suspension were added, and the plates were shaken at 23 °C for 90 min. The membranes were collected on GF/C filter mats (0.26 mm; Whatman) (pre-treated with 0.3% polyethylenimine for 30 min) and washed with cold Tris buffer (91 g l^{-1} Tris base, 25.5 g l^{-1} MgCl$_2$·6H$_2$O and 3.76 g l^{-1} EDTA) using a Brandel Harvester (Brandel). Filter pieces were punched out and transferred into 1450-401 96-well plates (PerkinElmer). Rotiscint eco plus (Carl Roth) (200 μl) was added, and the plates were sealed with transparent tape (permanent seal for microplates, PerkinElmer), vigorously shaken for at least 3 h and kept in the dark for at least 1 h before the measurement of radioactivity (d.p.m.) with a MicroBeta2 plate counter (PerkinElmer). Specific binding data (d.p.m.) were plotted against the free radioligand nanomolar concentration (obtained by subtracting the amount of bound radioligand (nM) (calculated from the specifically bound radioligand in d.p.m., the specific activity, and the volume per well) from the total radioligand concentration (nM)) and analysed by a two-parameter equation describing hyperbolic binding (SigmaPlot 11.0, Systat Software Inc.) to obtain K_d and receptor density (B_{max}) values. For K_d values < 1 nM, the B_{max} was kept below 1,200 d.p.m. by choosing an appropriate protein concentration. For K_d values > 1 nM, the B_{max} was kept below 3,300 d.p.m.

Competition-binding experiments. Competition-binding experiments were performed according to the procedure for saturation binding with the following modifications: [^{3}H]UR-MK299 was used undiluted and in the case of Y$_1$R mutants, for which [^{3}H]UR-MK299 exhibited a K_d value > 3 nM, the total volume per well was 100 μl, that is, in the case of total binding, wells were pre-filled with binding buffer (80 μl), and 10 μl of binding buffer containing [^{3}H]UR-MK299 (tenfold concentrated), and the membrane homogenate (10 μl) were added. The following concentrations of [^{3}H]UR-MK299 were used for competition binding with antagonists: 0.2 nM (wild-type Y$_1$R, T280$^{6.52}$A, T212$^{5.39}$A), 0.3 nM (F173$^{4.60}$W), 1.1 nM (L279$^{6.51}$A), 5 nM (Q219$^{5.46}$A), 7 nM (L215$^{5.42}$G), 10 nM (I124$^{3.36}$A, F173$^{4.60}$A). 1 nM [^{3}H]UR-MK299 was used for competition binding with NPY. The incubation time throughout was 90 min. Unspecific binding was determined in the presence of UR-MK299 (100-fold excess to [^{3}H]UR-MK299). Total binding was between 700 and 3,500 d.p.m. Maximum unspecific binding amounted to ≤30% of total binding. Specific binding data (d.p.m.) were plotted against log(concentration of competitor) and analysed by a four-parameter logistic equation (log(inhibitor) versus response–variable slope, GraphPad Prism Software 5.0) to obtain pIC$_{50}$ values, which were converted to half-maximal inhibitory concentration (IC$_{50}$) values (pIC$_{50}$ = $-\log_{10}$(IC$_{50}$)). In case of incomplete displacement of [^{3}H]UR-MK299 (specifically bound radioligand at the highest competitor concentration between 20% and 50%), pIC$_{50}$ values were determined by plotting log($B/(B_0 - B)$) (Hill plot; B denotes specifically bound radioligand in the presence of competitor (values between 10 and 90%), B_0 is the specifically bound radioligand in the absence of competitor ($B_0 = 100$%)) against log(competitor concentration) (at least three data points) and pIC$_{50}$ values (log($B/(B_0 - B)$) = 0) were determined by linear regression. K_i values were calculated from the IC$_{50}$ value as well as the respective K_d value (Extended Data Table 3) and the concentration of [^{3}H]UR-MK299 according to the Cheng–Prusoff equation[39].

Inositol phosphate accumulation assay. The signal transduction assay was performed as previously described[40,41]. In brief, COS-7 cells (obtained from American Type Culture Collection, and species specificity verified by COI DNA barcoding) were seeded into 48-well plates and were transiently co-transfected with wild-type receptor or receptor mutant and a chimeric G protein (G$_{\alpha\Delta6qi4myr}$) plasmid DNA[42]. Cells were routinely tested for mycoplasma contamination. Cells were radioactively labelled with myo[2-^{3}H]inositol (Perkin Elmer) overnight, and then stimulated with an increased concentration of NPY (NPY curve). For antagonist curves, cells were stimulated with the antagonist (BIBP3226: 10^{-5} M, BIBO3304: 10^{-6} M, MK-HU404: 10^{-7} M, UR-MK289: 10^{-5} M, UR-MK299: 10^{-7} M) parallel to an increased concentration of NPY for 1 h (standard conditions). After cell lysis, anion exchange chromatography was performed and isolated, radioactive accumulated inositol phosphate derivatives were analysed by liquid scintillation counting (scintillation cocktail Optiphase HiSafe, Perkin Elmer).

Using GraphPad Prism 5.0 (GraphPad Software) the determined concentration response curves were analysed. The curves were normalized to the top (100%) and bottom (0%) values of the associated NPY curve. All curves of independent experiments were summarized to one single concentration response curve by the row means total function. Using nonlinear regression (curve fit) the EC$_{50}$ and pEC$_{50}$ ± s.e.m. were examined. The shift between the NPY and NPY/antagonist curves is defined as the EC$_{50}$ ratio and calculated by dividing EC$_{50(NPY/antagonist)}$/EC$_{50(NPY)}$, the Hill slope was set to 1. All experiments were performed at least two times independently in technical duplicate.

Live-cell fluorescence microscopy. The membrane localization of Y$_1$R and receptor mutants was verified by fluorescence microscopy. COS-7 cells were seeded in 8-well μ-slides (IBIDItreat) and transiently transfected with Lipofectamine 2000 transfection reagent (Invivogen, Toulouse, France). Twenty-four hours after transfection, nuclei were stained with Hoechst33342 (Sigma-Aldrich) and images were recorded using an ApoTome Imaging System, AxioVert Observer Z1 (YFP: Filter Set 46, DAPI: Filter Set 49, ApoTome, 63×/1.40 oil objective, ZEISS) in a quasi-confocal setting. The data demonstrate that all the mutants are expressed at similar, and wild-type-like, level in COS-7 cells (Extended Data Fig. 2).

Quantification of receptor surface expression in COS-7 cells. COS-7 cells were seeded into black 96-well plates (Greiner), and transiently transfected with a plasmid encoding a receptor–eYFP fusion protein using MetafectenePro. Twenty-four hours after transfection, cells were washed once with HBSS, and fluorescence was quantified using a plate reader (Tecan Infinite M200, Tecan, Männedorf, Switzerland) at excitation 488/5 nm and emission 530/5 nm. Data were normalized to mock transfected (0%) and wild-type Y$_1$R–eYFP (100%). Results represent mean ± s.e.m. from three independent experiments performed in quadruplicate.

Peptide synthesis. Porcine NPY (YPSKPDNPGEDAPAEDLARYYSALRHYINLITRQRY–NH$_2$) and NPY analogues were synthesized by automated solid-phase peptide synthesis on an automated multiple peptide synthesis robot system (Syro, MultiSynTech), using a 9-fluorenylmethoxycarbonyl-*tert*-butyl (Fmoc/tBu) strategy in 15 μM scale as previously described[43]. NPY–tyramide was synthesized as previously described[44]. Isotopically labelled NPY variants were prepared as described[16], and ^{13}C/^{15}N-labelled amino acids were coupled manually with 2 equivalents (equiv.) hydroxybenzotriazole/ N,N′-diisopropylcarbodiimide (DIC) in DMF overnight. The porcine variant of NPY, which contains a single mutation M17L was used. This variant has binding affinity and signalling properties that are identical to human NPY and will therefore be referred to as wild-type NPY[45]. It also has increased solubility to assist in handling.

Modified NPY analogues [Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY and [K^4[(Ahx)$_2$-biotin]]NPY were synthesized by automated solid-phase peptide synthesis and Bpa, Ahx and biotin were coupled manually using orthogonal 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (Dde) protection groups, cleaved by freshly prepared 3% (v/v) hydrazine in DMF for 10 × 5 min. Manual coupling reactions were performed by incubation of the resin with 5 equiv. of the respective amino acid, 5 equiv. HOBt and 5 equiv. DIC in DMF for 2 h.

For biotin labelling, 3 equiv. biotin was dissolved in DMF for 10 min at 60 °C. Next, 3 equiv. HOBt and 3 equiv. DIC were added to the mixture. Coupling was performed overnight at room temperature under constant shaking. Bpa containing peptides were cleaved from the resin and completely deprotected with a mixture of trifluoroacetic acid (TFA)/thioanisole (TA)/water (90:5:5 v/v/v).

All peptides were purified by preparative reversed-phase high-performance liquid chromatography (RP-HPLC) on a Jupiter 4 u Proteo RP-C18 column (90 Å, 4 μm, Phenomenex), Kinetex 5 u XB-C18 column (100 Å, 5 μm, Phenomenex), Kinetex 5 u Biphenyl (100 Å, 5 μm, Phenomenex), Aeris 3.6 μm WIDEPORE XB-C18 (200 Å, 3.6 μm, Phenomenex) or Varitide RPC (200 Å, 6 μm, Agilent Technologies). All peptides were characterized by matrix-assisted laser desorption/ionization time of flight (MALDI–TOF) mass spectrometry (Ultraflex III MALDI–TOF/TOF, Bruker daltonics) and ESI-HCT (high-capacity ion trap electrospray-ionization mass spectrometry, Bruker Daltonics). Peptide purities were determined by two different analytical RP-HPLC systems using 0.1% (v/v) TFA in H$_2$O (eluant A) and 0.08% (v/v) TFA in acetonitrile (ACN) (eluant B). Purity of all peptides was ≥ 95%.

NMR measurements of Y$_1$R-bound NPY. Fourteen differently isotopically labelled [U-^{13}C/^{15}N]porcine NPY peptides were prepared by standard fluorenylmethyloxycarbonyl (Fmoc) solid-phase synthesis as described previously[16]. On the basis of the structure of micelle-bound NPY, the positions of the NMR labels were chosen to avoid signal overlap in ^{13}C–^{13}C single quantum/double quantum correlation experiments and to allow straightforward signal assignment.

Expression of the human Y$_1$R in *Escherichia coli* as inclusion bodies, inclusion body preparation, solubilization of the receptor in SDS and receptor purification were as described[46] yielding ~40–50 mg Y$_1$R per litre of expression medium. To assemble the Y$_1$R into a functional state, a three-step folding protocol was applied. In step 1, the purified Y$_1$ receptor was dialysed against a degassed buffer containing 2 mM SDS, 50 mM NaP, pH 8.5, 1 mM EDTA, 1 mM reduced glutathione, and

0.5 mM oxidized glutathione at room temperature for 48 h. Subsequently, 25% (w/w) poly(ethylene glycol) (molecular mass 20 kDa) was added to the buffer to concentrate the receptor before reconstitution. In step 2, bicelles consisting of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC-c7) (AvantiPolarLipids) with a *q* value of 0.25 in 50 mM NaP, pH 8.0 were incubated with Y_1R, with three subsequent cycles of fast temperature cycles from 42 °C to 0 °C. Aggregated protein was removed instantly by centrifugation. In step 3, the Y_1R samples were concentrated in large bicelles ($q > 20$), which were used instead of liposomes because of the high achievable receptor packing[46]. In large bicelles, all receptor binding sites are fully accessible. Subsequently, 50 mg ml^{-1} BioBeadsSM2 was added at least twice to the solution. After removal of the beads with a sieve, samples were washed four times through cycles of pelleting by centrifugation and resolubilization in 50 mM NaP, pH 7.0. Concentration determination of the membrane embedded receptor was performed by solubilization of the bicelles in ten times the volume of 15 mM SDS and 50 mM NaP, pH 7.0 and subsequent measurement of the Y_1R intrinsic absorption at 280 nm using UV–visible spectroscopy. Labelled NPY variants in a slight molar excess were added to the Y_1R after detergent removal but before concentrating.

Assessment of the binding affinity of the Y_1R was carried out using homogenous fluorescence assays as described in the literature[47]. The reconstituted receptor was incubated in increasing concentrations with 50 nM fluorescently labelled NPY (NPY-atto520) overnight at room temperature in 50 mM NaP, pH 7.0 and 0.1% BSA. The fluorescence spectra were recorded on a FluoroMax-2 (JOBIN YVON) in a 10 mm quartz cuvette at 20 °C. The maximum signals of each spectrum were determined, normalized and plotted against the receptor concentration. The inflection point for Y_1R binding was determined (OriginPro 8 G / DoseResp) at $EC_{50} = 52$ nM, demonstrating high functionality of the system. As a control, we used empty bicelles in concentrations that matched the bicelle concentration of the receptor samples, resulting in a lower binding ability to the ligand in comparison to the Y_1R.

NMR spectra were recorded on a Bruker Avance III 600 NMR spectrometer. The ^{13}C cross-polarization magic angle spinning (MAS) NMR experiments (0.7 ms contact time) were carried out using a double-resonance MAS probe with a 4 mm spinning module. Typical 90° pulse lengths were 4 µs for ^1H and ^{13}C and heteronuclear decoupling (SPINAL64) at a field strength of ~65 kHz. The ^{13}C chemical shifts were referenced relative to tetramethylsilane. The experiments were conducted at −30 °C and an MAS frequency of 7 kHz. The ^{13}C double quantum ^{13}C single quantum correlation spectra were acquired using the SPC-5 recoupling sequence[48] for double quantum excitation and reconversion (set to 0.571 ms each). The relaxation delay was 2.4 s.

Molecular docking of NPY into Y_1R. Peptide docking of full-length porcine NPY was completed using Rosetta's FlexPepDock application[49]. In brief, low energy backbone conformations were generated from the starting conformation of UR-MK299-bound Y_1R. Initially, the trimer of C-terminal NPY was docked into these conformations using full flexible docking guided by mutagenesis data. For each round of docking 5,000 models were generated. The models were sorted by total energy and binding energy. Top models from a given docking round were used to seed the next round of docking in which the peptide was extended. Fragment picking was performed using the fragment_picker application within Rosetta[26]. Secondary structure during fragment picking was guided by the NMR chemical shift data. Additionally, experimentally derived restraints were used to guide docking (R35–D287[6.59], R35–N283[6.55], Y36–Y100[2.64], R33–N299[7.32], L30–I293[ECL3]). After docking peptides of length 6, 12, 18 and 36, the binding pocket was resampled to allow the ligand binding pocket to adapt to the shape of the peptide. This was accomplished with RosettaCM[50]. The Y_1R crystal structure was used as a template along with the docked model to ensure the models did not drift too far from the starting structure though the N terminus was removed until the last docking step to provide steric bulk. Following full-length NPY docking, the N terminus of NPY was localized using loose distance constraints with the peptides identified in crosslinking experiments. Model selection from RosettaCM was accomplished using clustering to ensure backbone diversity. Following the identification of docked poses that satisfied the majority of experimental data, the chemical shifts of docked NPY peptides were calculated and filtered against the experimental NMR data to generate a final ensemble of docked poses with a 1.4 p.p.m. root mean square distance to the experimental data. To analyse the binding interactions, per residue energetic analysis was calculated using the residue_energy_breakdown algorithm. The model with the lowest energy was selected as the representative binding pose (Fig. 4a). The ensemble is rather tight and therefore the individual binding poses are similar in overall structure (Fig. 4b).

Photo-crosslinking experiment between Y_1R and NPY N terminus. Cell-free expressed Y_1R was produced by coupled in vitro transcription–translation performed as previously described[51] using a bacterial cell lysate (S30 extract) from *E. coli* BL21 (DE3). Soluble membrane protein expression was achieved by

addition of 0.1% (w/v) Brij-58, 1 mM oxidized glutathione (GSSG) and 5 mM reduced glutathione (GSH). Expression buffer was then exchanged to a binding buffer (0.1 M Tris/HCl, pH 7.4, 5% glycerol and 0.1% (w/v) Brij-58) and samples were purified by ligand affinity chromatography using [K^4[(Ahx)$_2$-biotin]]NPY immobilized on Pierce Avidin Agarose beads (Thermo Fisher Scientific) as previously described[52]. Elution was performed using 60 mM CaCl$_2$.

For photo-crosslinking Y_1R in binding buffer was incubated with [Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY in a molar ratio of 4:1 (5 nmol:1.25 nmol) for 30 min at room temperature. In addition, the same reaction was performed with an eightfold-excess of NPY (Y_1R:[Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY:NPY, 4:1:8). Subsequently, the opened reaction vessels were placed on ice and irradiated with UV light (UV lamp, Atkas Fluorest forte, $\lambda = 366$ nm, 180 W) for 90 min. 50 µl of photo-crosslinked Y_1R sample (~200 µg) was digested with Glu-C and rLys-C (Promega) according to the manufacturer's protocol. Crosslinked fragments were then isolated by affinity purification using Monomeric Avidin Agarose beads (Thermo Fisher Scientific) according to the manufacturer's protocol. Possible fragments of digested Y_1R were calculated using the online tool PeptideMass[53]. To account for incomplete digestion the tool allowed for a maximum of five missed cleavages. For the analysis the combined option 'Glu C (phosphate) + Lys-C' was chosen. The same procedure was used for the calculation of possible NPY fragments. Theoretical masses of fragments after enzymatic cleavage of photo-crosslinked Y_1R–[Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY were calculated by adding possible Y_1R fragment masses to NPY fragment masses containing the N terminus. The masses of Bpa, two times Ahx and biotin reduced by water were added manually to account for the formation of a peptide bond. Peptide fragments of photo-crosslinked Y_1R were analysed by MALDI–TOF mass spectrometry using an Ultraflex III MALDI–TOF/TOF mass spectrometer (Bruker Daltonics).

Functionality of cell-free expressed Y_1R samples was verified by a homogenous binding assay based on fluorescence polarization. We used [Dpr22-Atto 520]NPY (hereafter: NPY-Atto 520) as a fluorescence tracer (inositol phosphate accumulation in transiently transfected COS-7: $EC_{50} = 24$ nM, $pEC_{50} = -7.61 \pm 0.20$). 50 nM of NPY-Atto 520 was incubated with increasing concentrations of Y_1R in Brij-58 micelles in buffer (0.1 M Tris/HCl, pH 7.4, 2.5% glycerol, 0.1% (w/v) Brij-58 and 0.1% bovine serum album) for 90 min under gentle agitation in opaque 96-well plates. Fluorescence was then measured in a Tecan Spark plate reader (Tecan) using linear polarized light (excitation 510/5 nm, emission 550/10 nm, 90° detection angle). Experiments were conducted at least twice independently in duplicate.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. Atomic coordinates and structure factor files for the Y_1R–UR-MK299 and Y_1R–BMS-193885 complex structures have been deposited in the Protein Data Bank (PDB) with accession codes 5ZBQ and 5ZBH, respectively.

31. Rosenbaum, D. M. et al. GPCR engineering yields high-resolution structural insights into β$_2$-adrenergic receptor function. *Science* **318**, 1266–1273 (2007).
32. Roth, C. B., Hanson, M. A. & Stevens, R. C. Stabilization of the human β$_2$-adrenergic receptor TM4–TM3–TM5 helix interface by mutagenesis of Glu122[3.41], a critical residue in GPCR structure. *J. Mol. Biol.* **376**, 1305–1319 (2008).
33. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
34. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
35. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D.* **67**, 355–367 (2011).
36. Smart, O. S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D.* **68**, 368–380 (2012).
37. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D.* **66**, 486–501 (2010).
38. Keller, M. et al. Mimicking of arginine by functionalized N$^\omega$-carbamoylated arginine as a new broadly applicable approach to labeled bioactive peptides: high affinity angiotensin, neuropeptide Y, neuropeptide FF, and neurotensin receptor ligands as examples. *J. Med. Chem.* **59**, 1925–1945 (2016).
39. Yung-Chi, C. & Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I_{50}) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
40. Burkert, K. et al. A deep hydrophobic binding cavity is the main interaction for different Y$_2$R antagonists. *ChemMedChem.* **12**, 75–85 (2017).
41. Els, S., Beck-Sickinger, A. G. & Chollet, C. Ghrelin receptor: high constitutive activity and methods for developing inverse agonists. *Methods Enzymol.* **485**, 103–121 (2010).
42. Kostenis, E. Is Gα$_{16}$ the optimal tool for fishing ligands of orphan G-protein-coupled receptors? *Trends Pharmacol. Sci.* **22**, 560–564 (2001).
43. Pedragosa-Badia, X. et al. Pancreatic polypeptide is recognized by two hydrophobic domains of the human Y$_4$ receptor binding pocket. *J. Biol. Chem.* **289**, 5846–5859 (2014).
44. Hoffmann, S., Rist, B., Videnov, G., Jung, G. & Beck-Sickinger, A. G. Structure-affinity studies of C-terminally modified analogs of neuropeptide Y led to a novel class of peptidic Y$_1$ receptor antagonist. *Regul. Pept.* **65**, 61–70 (1996).

45. Gerald, C. et al. A receptor subtype involved in neuropeptide-Y-induced food intake. *Nature* **382**, 168–171 (1996).
46. Schmidt, P. et al. A reconstitution protocol for the in vitro folded human G protein-coupled Y_2 receptor into lipid environment. *Biophys. Chem.* **150**, 29–36 (2010).
47. Casiraghi, M. et al. Functional modulation of a G protein-coupled receptor conformational landscape in a lipid bilayer. *J. Am. Chem. Soc.* **138**, 11170–11175 (2016).
48. Hohwy, M., Rienstra, C. M., Jaroniec, C. P. & Griffin, R. G. Fivefold symmetric homonuclear dipolar recoupling in rotating solids: application to double quantum spectroscopy. *J. Chem. Phys.* **110**, 7983–7992 (1999).
49. Raveh, B., London, N., Zimmerman, L. & Schueler-Furman, O. Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS ONE* **6**, e18934 (2011).
50. Song, Y. et al. High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742 (2013).
51. Schwarz, D. et al. Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. *Nat. Protocols* **2**, 2945–2957 (2007).
52. Bosse, M. et al. Assessment of a fully active class A G protein-coupled receptor isolated from in vitro folding. *Biochemistry* **50**, 9817–9825 (2011).
53. Wilkins, M. R. et al. Detailed peptide characterization using PEPTIDEMASS—a World-Wide-Web-accessible tool. *Electrophoresis* **18**, 403–408 (1997).
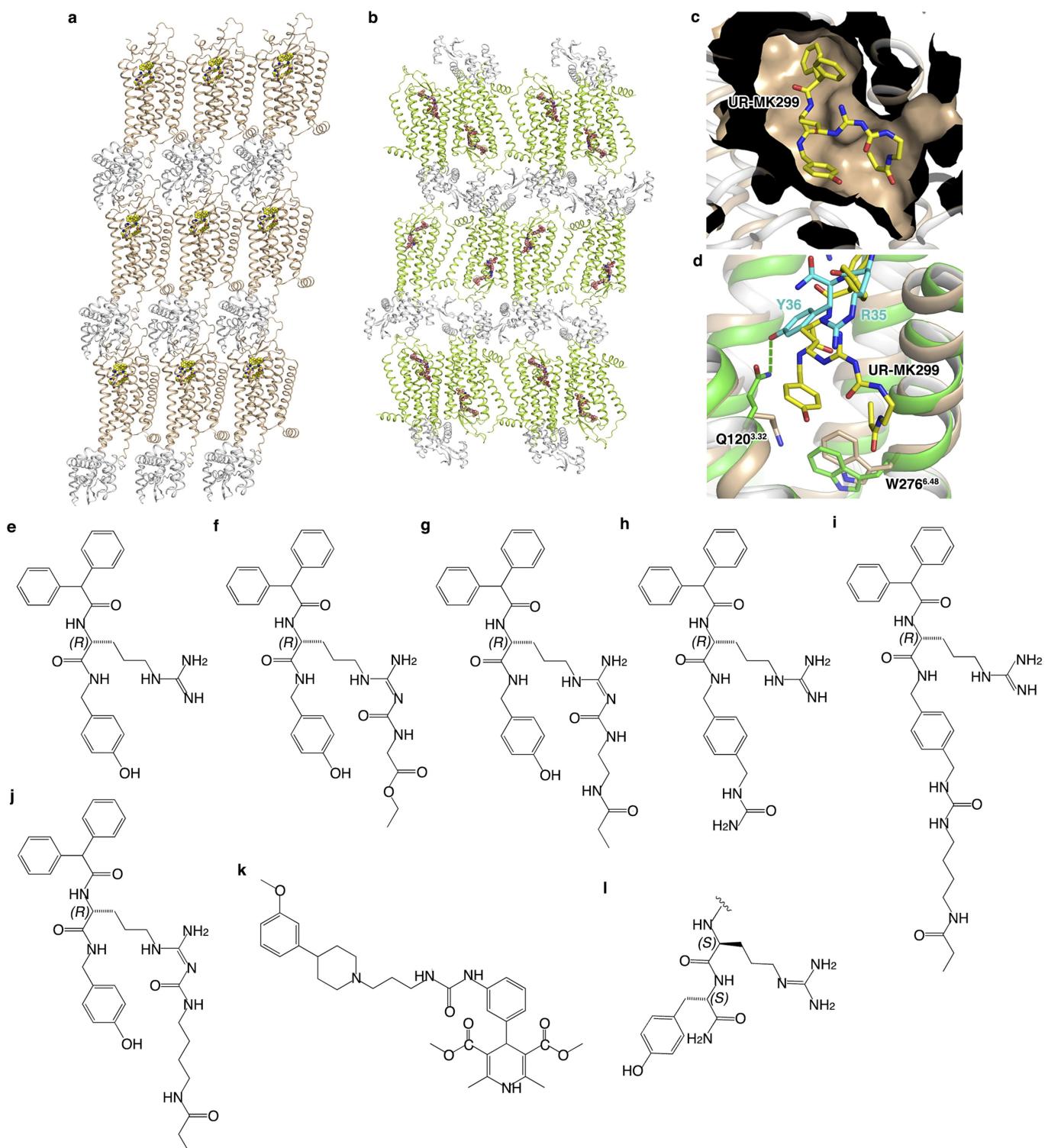
Extended Data Fig. 1 | Crystal packing and structural features of Y_1R and chemical structures of Y_1R ligands. a, b, Crystal packing of Y_1R–UR-MK299 (**a**) and Y_1R–BMS-193885 (**b**) complexes. Y_1R is shown in cartoon representation and coloured brown and green in the Y_1R–UR-MK299 and Y_1R–BMS-193885 complexes, respectively. The T4L fusion is shown in grey cartoon representation. UR-MK299 and BMS-193885 are displayed as yellow and pink spheres, respectively. **c,** Cutaway view of the UR-MK299-binding pocket in Y_1R. The receptor is shown in brown cartoon and surface representations. The ligand is shown as yellow sticks. **d,** Comparison of Y_1R in the Y_1R–UR-MK299 crystal structure (brown)

and the Y_1R–NPY model (green). Side chains of Q120$^{3.32}$ and W276$^{6.48}$ are shown as sticks. R35–Y36 of NPY is displayed as cyan sticks. The hydrogen bond between Q120$^{3.32}$ and Y36 of NPY is shown as a green dashed line. **e–j,** Chemical structures of the argininamide Y_1R antagonists BIBP3226 (**e**), UR-HU404 (**f**), UR-MK299 (**g**), BIBO3304 (**h**), UR-MK289 (**i**) and UR-MK136 (**j**). **k,** Chemical structure of BMS-193885. **l,** Scaffold of NPY C-terminal residues R35 and Y36. Key differences between R35–Y36 of NPY and UR-MK299 are chirality of the arginine derivative and alteration of bond connectivity leading to the hydroxyphenyl group.
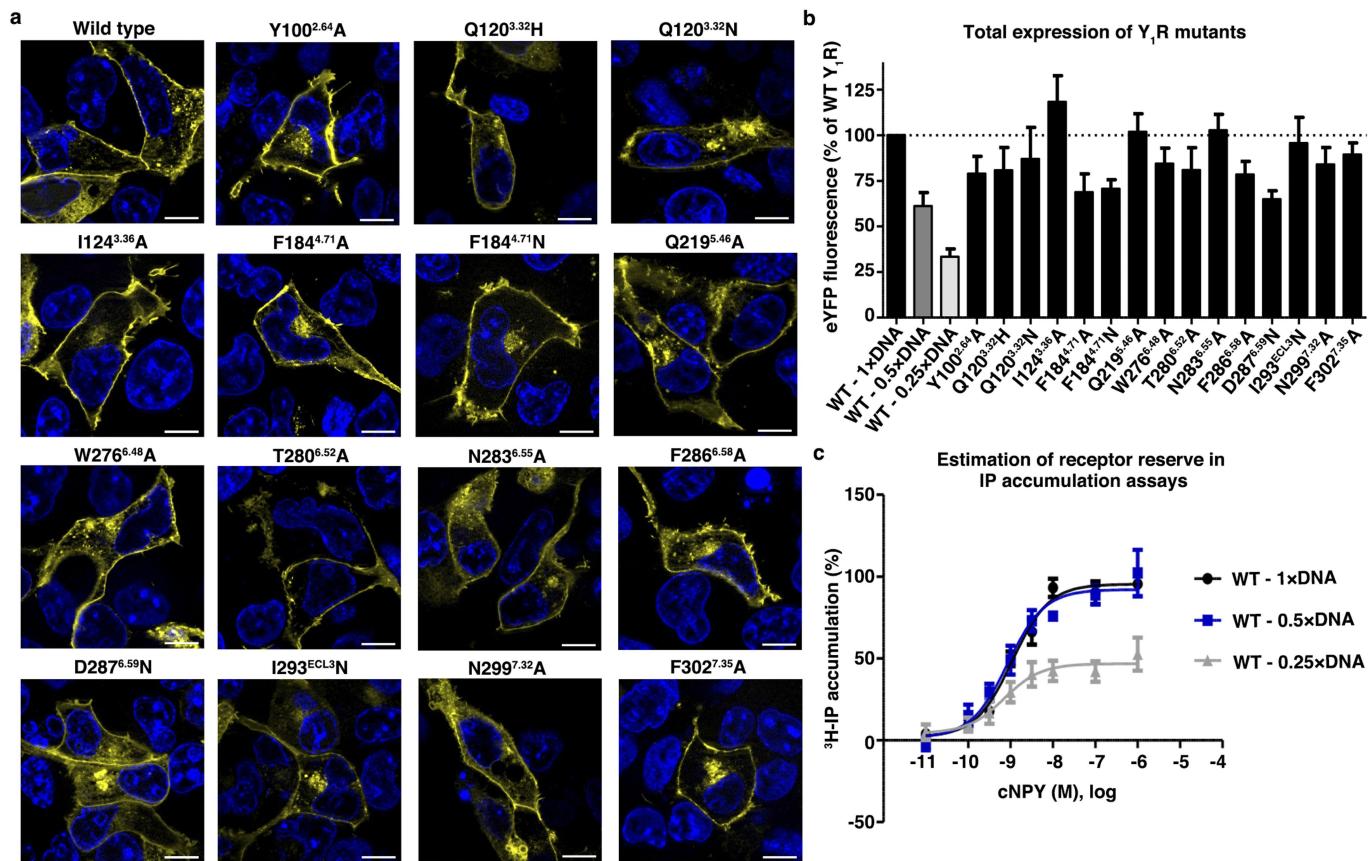
Extended Data Fig. 2 | Expression of wild-type and mutant Y_1 receptors in transiently transfected COS-7 cells. a, Live-cell fluorescence microscopy verifies all Y_1R variants to be properly folded and exported to the cell membrane like the wild-type receptor. Nuclei stained with Hoechst33342. Scale bars, 10 μm. Pictures are representative of two independent experiments with similar results. **b**, The total expression level was determined by fluorescence reading and expression was confirmed to be similar to the wild type. Transfection of only 50% or 25% of the DNA amount (with total DNA amount held constant by empty vector), led to a proportional decrease of fluorescence, and thus, expression level. Data represent mean ± s.e.m. of three to five independent experiments performed in technical triplicate (see Source Data for sample size of each mutant). **c**, Estimation of the receptor reserve in functional inositol phosphate accumulation assays. Transfection of half of the vector encoding the receptor (with a constant total DNA amount including chimeric G protein, see **a**) still produces maximum signal, while further reduction results in signal loss at comparable potency. Thus, there is only a small receptor reserve in the functional readout, allowing potency alteration to be directly related to compromised ligand binding. Data represent mean ± s.e.m. of three independent experiments performed in technical duplicate. cNPY, concentration of NPY.

```
NPY1R_HUMAN  ...........................................................................
NPY2R_HUMAN  ....................................................................MGPIG
NPY4R_HUMAN  ...........................................................................
NPY5R_HUMAN  ...........................................................................
NPFF1_HUMAN  ...........................................................................
NPFF2_HUMAN  MNSFFGTPAASWCLLESDVSSAPDKEAGRERRALSVQQRGGPAWSGSLEWSRQSAGDRRRLGLSRQTAKSSWSRSRDRTCCCRRAWWILVPAA
```

```
                      1        10                  20        30        40        50        60        70
NPY1R_HUMAN  ....MNST.LFSQVENH...........SVHSNFSEKNAQLLAF.ENDDCHLPLAMIFTLALAYGAVIILGVSGNLALIIIILKQKEMRNVTN
NPY2R_HUMAN  AEADENQTVEEMKVEQY...........GPQTTPRGELVPDPEP.ELIDSTKLIEVQVVLILAYCSILLLGVIGNSLVIHVVIKFKSMRTVTN
NPY4R_HUMAN  ....MNTSHLLALLLPK..........SPQGENRSKPLGTPYN.FSEHCQDSVDVMVFIVTSYSIETVVGVLGNLCLMCVTVRQKEKANVTN
NPY5R_HUMAN  ..........MDLELDEYYNKTL.....ATENNTAATRNSDFPVWDDYKSSVDDLQYFLIGLYTFVSLLGFMGNLLILMALMKKRNQKTTVN
NPFF1_HUMAN  ..........MEGEPSQPPNSSWPL..SQNGTNTEATPATNLT.FSSYYQHTSPVAAMFIVAYALIFLLCMVGNTLVCFIVLKNRHMHTVTN
NPFF2_HUMAN  DRARRE..RFIMNEKWDTNSSENWHPIWNVNDTKHHLYSDINIT.YVNYYLHQPQVAAIFIISYFLIFFLCMMGNTVVCFIVMRNKHMHTVTN
```

```
                 80        90        100       110       120       130       140       150       160
NPY1R_HUMAN  ILIVNLSFSDLLVAIMCLPFTFVYTLMDHWVFGEAMCKLNPFVQCVSITVSIFSLVLIAVERHQLIINPRGWRPNNRHAYVGIAVIWVLAVAS
NPY2R_HUMAN  FFIANLAVADLLVNTLCLPFTLTYTLMGEWKMGPVLCHLVPYAQGLAVQVSTITLTVIALDRHRCIVYHLESKISKRISFLIIGLAWGISALL
NPY4R_HUMAN  LLIANLAFSDFLMCLLCQPLTAVYTIMDYWIFGETLCKMSAFIQCMSVTVSILSLVLVALERHQLIINPTGWKPSISQAYLGIVLIWVIACVL
NPY5R_HUMAN  FLIGNLAFSDILVLVLFCSPFTLTSVLLDQWMFGKVMCHIMPFLQCVSVLVSTLILISIAIVRYHMIKHPISNNLTANHGYFLIATVWTLGFAI
NPFF1_HUMAN  MFILNLAVSDLLVGIFCMPTTLVDNLITGWPFDNATCKMSGLVQGMSVSASVFTLVAIAVERFRCIVHPFREKLTLRKALVTIAVIWALALLI
NPFF2_HUMAN  LFILNLAISDLLVGIFCMPITLLDNIIAGWPFGNTMCKISGLVQGISVAASVFTLVAIAVDRFQCMVVYPFKPKLTIKTAFVIIMIIWVLAITI
```

F4.60 ↓   Q3.32   I3.36   Q5.46 ↓

```
                170       180       190       200       210       220       230       240
NPY1R_HUMAN  SLPFLIYQVMTDEPF...QNVTLDAYKDKYVCFDQFPSDSHRL...SYTTLLLVLQYFGPLCFIFICYFKIYIRLKRR...............
NPY2R_HUMAN  ASPLAIFREYSL.......IEIIPDFEIVACTEKWPGEEKSIYGTVYSLSSLLILYVLPLGIISFSYTRIWSKLKNH...............
NPY4R_HUMAN  SLPFLANSILENVFH.KNHSKALEFLADKVVCTESWPLAHHRT...IYTFTLLLFQYCLPLGFILVCYARIYRRLQRQ...............
NPY5R_HUMAN  CSPLPVFHSLVELQE....TFGSALLSSRYLCVESWPSDSYRI...AFTISLLLVQYILPLVCLTVSHTSVCRSISCGLSNKENRLEENEMIN
NPFF1_HUMAN  MCPSAVTLTVTREEH.HFMVDARNRSYPLYSCWEAWPEKGMRR...VYTTVLFSHIYLAPLALIVMMYARIARKLCQA...............
NPFF2_HUMAN  MSPSAVMLHVQEEKYYRVRLNSQNKTSPVYWCREDWPNQEMRK...IYTTVLFANIYLAPLSLIVIMYGRIGISLFRA...............
```

T5.39 ↑

```
                                  250
NPY1R_HUMAN  ........NNMMDKM...RDNK...............
NPY2R_HUMAN  .........VSPGAAN...DHYH..............
NPY4R_HUMAN  ..........GRVFHKG...TYS..............
NPY5R_HUMAN  LTLHPSKKSGPQVKLSGGSHKWSYSFIKKHRRRYSKKTACVLPAPERPSQENHSRILPENFGSVRSQLSSSSKFIPGVPTCFEIKPEENSDVHE
NPFF1_HUMAN  .........PGPAPGGEEAAD.P..............
NPFF2_HUMAN  .........AVPHTGRKNQEQWH..............
```

N6.55 ↓   F6.58 ↓

```
                260       270       280       290       300       310       320       330
NPY1R_HUMAN  ........YRSSETKRINIMLLSIVVAFAVCWLPLTIFNTVFDWNHQIIATCN..HNLLFLLCHLTAMISTCVNPIFYGFLNKNFQRDLQFFF
NPY2R_HUMAN  .......QRRQKTTKMLVCVVVVFAVSWLPLHAFQLAVDIDSQVLDLKE..YKLIFTVFHIIAMCSTFANPLLYGWMNSNYRKAFLSAF
NPY4R_HUMAN  .......LRAGHMKQVNNVVLVMVVVFAVLWLPLHVFNSLEDWHHEAIPICH..GNLIFLVCHLLAMASTCVNPFIYGFLNTNFKKEIKALV
NPY5R_HUMAN  LRVKRSVTRIKKRSRSVFYRLTILILVFAVSWMPLHLFHVVTDFNDNLISNRH..FKLVYCICHLLGMMSCCLNPILYGFLNNGIKADLVSLI
NPFF1_HUMAN  .......RASRRRARVHHMLVMVALFFTLSWLPIWALLLLIDYGQLSAPQLHLVTVYAFPFAHWLAFFNSSANPIIYGYFNENFRRGFQAAF
NPFF2_HUMAN  .......VVSRKKQKIIKMLLIVALLFILSWLPLWTLMMLSDYADLSPNELQIINIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRGFQEAF
```

W6.48 ↑   L6.51 ↑   F6.54 ↑   D6.59 ↑   F7.35 ↑

```
                340       350       360       370       380
NPY1R_HUMAN  N..FCDFRSRDDD..YETIAMST.MHTDVSKT.....SLKQASPVAFKKINNNDDNEK.....I...............
NPY2R_HUMAN  R...CEQRLD...........A.IHSEVSVTFKAKKNLEVRKNSGPNDSFTEATNV..............
NPY4R_HUMAN  L..TCQQSAPLEE..SEHLPLST.VHTEVSKG....SLRLSGRSNPI..............
NPY5R_HUMAN  H...CLHM...............
NPFF1_HUMAN  RARLCPRPSGSHKEAYSERPG.GLLHRRVFVVV.RPSDSGLPSESGPSSGAPRPGRLPLRNGRVAHHGLPREGPGCSHLPLTIPAWDI
NPFF2_HUMAN  QLQLCQKRAKPME.AYALKAKSHVLINTSNQLV.QEST..FQNPHGETLLYRKSAEKPQQE..LVMEEL.KETTNSSEI........
```

Extended Data Fig. 3 | Sequence alignment of the human NPY receptors and the human NPFF receptors. Colours represent the similarity of residues: red background, identical; red text, strongly similar. Key residues in the UR-MK299-binding pocket, which are conserved or variable among receptors, are indicated by red or black arrows, respectively. The alignment was generated using UniProt (http://www.uniprot.org/align/) and the graphic was prepared on the ESPript 3.0 server (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi).

a

b

c

	C_α	C_β	C_γ	C_δ	C_ε
Y1	57.2	39.0			
P2	62.0	31.7	24.6	48.0	
S2	58.0	63.3			
K4	53.1	32.3	25.1	28.8	41.0
P5	62.1	30.1	27.0	48.5	
D6	52.6	40.0			
N7	48.0	37.0			
P8	56.7	28.1	25.9	48.8	
G9					
E10	54.9	32.1	27.9		
D11	56.0	38.7			
A12	48.5	16.0			
P13	60.1	30.8	24.7	46.7	
A14	53.3	17.1			
E15	56.6	27.7	27.7		
L17	55.6	39.7	25.5	25.5	25.5
A18	53.3	16.7			
R19	57.8	29.2	24.7	41.5	
Y20	60.5	35.1			
Y21	60.3	36.3			
S22	60.8	60.1			
A23	52.9	16.6			
L24	56.9	39.8	25.5	25.5	25.5
R25	57.7	28.0	24.4	41.1	
Y27	60.2	37.4			
I28	63.5	36.2	28.1	16.6	12.4
L30	56.5	39.8	25.8	25.8	25.8
I31	63.0	36.7	27.4	16.8	12.7
T32	60.2	65.4	18.0		
R33	54.3	29.1	23.7	42.7	54.3
Q34	54.0	27.4	29.1		
R35	52.5	29.6	25.0	41.8	
Y36	54.7	38.7			

d

Extended Data Fig. 4 | Pharmacological characterization of refolded Y_1R and NMR studies of Y_1R-bound NPY. a, Binding of Atto 520-labelled NPY (50 nM) to increasing amounts of bicelles containing Y_1R or empty bicelles. Data reflect fluorescence enhancement upon binding. An inflection point at $EC_{50} = 52$ nM was determined. Two independent experiments were performed in technical duplicate with similar results. Data shown are from a representative experiment. a.u., arbitrary units. $c(Y_1R)$, concentration of Y_1R. **b**, Typical ^{13}C MAS single-quantum (SQ)/double-quantum (DQ) correlation spectrum of NPY in the presence of Y_1R reconstituted into large bicelles at $-30\,°C$. NMR spectra were acquired from one to three independent preparations for each labelled amino acid with similar results (see **d**). Data shown are from a representative experiment. **c**, Table showing ^{13}C-NMR chemical shifts of assigned amino acids of NPY bound to Y_1R (referenced to tetramethylsilane) as acquired in solid-state NMR experiments. **d**, ^{13}C-chemical-shift index of NPY bound to Y_1R in large DMPC/DHPC-c7 bicelles ($q > 20$) compared with docked models. Plotted in black is the measured chemical shift difference ($C_\alpha - C_\beta$) for each individual residue of NPY minus the chemical shift difference of the same amino acid type in random-coil conformation. Individual data points from one to three independent experiments for each labelled amino acid are shown. Typical experimental error when determining chemical shifts under these conditions are ± 1 p.p.m. Chemical shifts were back-calculated for the top docking solutions and filtered against the experimental data to generate a final ensemble of docked poses. Their average chemical-shift index and associated s.d. from the top ten docked poses are shown in red.
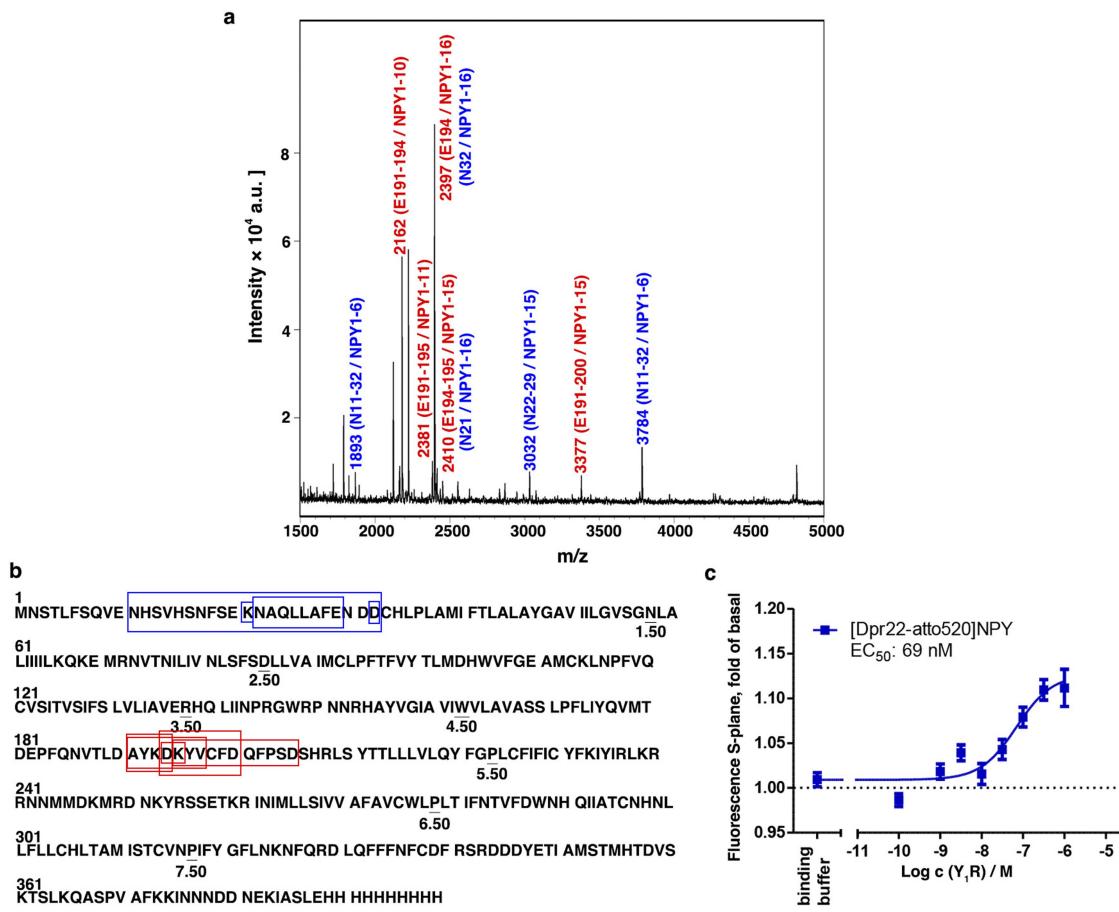
Extended Data Fig. 5 | Photo-crosslinking experiments between NPY and Y$_1$R. a, Mass spectra of photo-crosslinked Y$_1$R with [Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY. Exemplary MALDI–TOF mass spectra of photo-crosslinked samples enzymatically digested by rLys-C and Glu-C. Potential Y$_1$R fragments are labelled. Two independent experiments were performed with similar results. N, N terminus of Y$_1$R (blue); E, ECL2 (red). **b**, Respective regions of NPY N terminus at Y$_1$R. Amino acid sequence of Y$_1$R with a C-terminal His-tag. The two detected regions within Y$_1$R (N terminus (blue), ECL2 (red)) after crosslinking with [Bpa^1,K^4[(Ahx)$_2$-biotin]]NPY are emphasized in boxes. The different sizes of the boxes represent different detected fragments (Extended Data Table 5). Experiments were repeated twice independently with similar results, and only fragments that were observed in both experiments are listed here and in Extended Data Table 5. **c**, Binding of Atto 520-labelled NPY (50 nM) to increasing amounts of cell-free produced Y$_1$R in Brij-58. Data reflect fluorescence enhancement upon binding. An EC$_{50}$ value of 69 nM was determined. Data shown are mean ± s.e.m. from six independent experiments performed in technical triplicate. c(Y$_1$R), concentration of Y$_1$R.

Extended Data Table 1 | Data collection and refinement statistics

	Y_1R–UR-MK299	Y_1R–BMS-193885
Data Collection*		
Space group	$P2_1$	$C222_1$
Cell dimensions		
a, b, c (Å)	37.8, 100.7, 83.2	76.9, 126.8,170.3
α, β, γ (°)	90.0, 98.8, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50.0-2.70 (2.83-2.70)†	50.0-3.0 (3.1-3.0)
R_{merge} (%)	17.0 (86.1)	16.6 (93.7)
$I / \sigma I$	4.78 (0.97)	5.30 (1.00)
Completeness (%)	97.3 (96.9)	92.4 (79.2)
Redundancy	4.0 (3.6)	3.5 (2.6)
Refinement		
Resolution (Å)	50.0-2.7	50.0-3.0
No. reflections	16,520 (790)	15,600 (797)
R_{work} / R_{free} (%)	22.5 / 24.7	22.4 / 24.9
Number of atoms		
Protein	3,715	3,654
Ligand	45	43
Overall B values (Å²)		
Protein	88.2	108.0
Ligand	65.2	81.0
R.m.s. deviations		
Bond lengths (Å)	0.010	0.009
Bond angles (°)	1.04	1.00

*Diffraction data from 47 Y_1R–UR-MK299 crystals and 33 Y_1R–BMS-193885 crystals were used to solve the structures.
$\dagger$Numbers in parentheses refer to the highest-resolution shell.

Extended Data Table 2 | Inositol phosphate accumulation assays of wild-type and mutant Y₁ receptors for NPY and antagonists

Mutants	NPY EC₅₀ (nM)† (pEC₅₀ ± SEM)	n‡	NPY/BIBP3226 (10⁻⁵ M)* EC₅₀ (nM) (pEC₅₀ ± SEM)	Ratio§	K_b (nM)‖	n	NPY/BIBO3304 (10⁻⁶ M) EC₅₀ (nM) (pEC₅₀ ± SEM)	Ratio	K_b (nM)	n	NPY/UR-HU404 (10⁻⁷ M) EC₅₀ (nM) (pEC₅₀ ± SEM)	Ratio	K_b (nM)	n	NPY/UR-MK289 (10⁻⁵ M) EC₅₀ (nM) (pEC₅₀ ± SEM)	Ratio	K_b (nM)	n	NPY/UR-MK299 (10⁻⁷ M) EC₅₀ (nM) (pEC₅₀ ± SEM)	Ratio	K_b (nM)	n
Wild type	1.7 (8.78 ± 0.03)	20	463 (6.43 ±0.04)	272	36.9	14	175 (6.76 ± 0.04)	103	9.8	16	1,099/126# (5.96 ± 0.04)/(6.90 ± 0.07)	646/74#	0.16/0.14#	14	228 (6.64 ± 0.04)	134	75.2	15	250 (6.60 ± 0.06)	147	0.69	10
Y100²·⁶⁴A	562; 404	2	nd	nd	nd	2	nd	nd	nd	2	nd	nd	nd	2	nd	nd	nd	2	/	/	/	/
Q120³·³²H	12; 15	2	nd	nd	nd	2	nd	nd	nd	2	>50,000; >50,000	>3,500	<0.03	2	>50,000 >50,000	>3,500	<2.86	2	>50,000 >50,000	>3,500	< 0.03	2
Q120³·³²N	3.8 (8.42 ± 0.07)	4	836 (6.08 ± 0.06)	220	45.7	3	414 (6.38 ± 0.09)	109	9.3	3	3,744 (5.43 ± 0.10)	985	0.10	3	325 (6.49 ± 0.08)	86	118	3	743 (6.13 ± 0.06)	196	0.51	3
I124³·³⁶A	6.0 (8.22 ± 0.08)	3	1,877 (5.73 ± 0.06)	313	32.1	4	306 (6.52 ± 0.06)	51	20.0	4	1,588 (5.80 ± 0.06)	265	0.38	4	637 (6.20 ± 0.08)	106	95.2	4	208 (6.68 ± 0.09)	35	2.94	4
Q219⁵·⁴⁶A	23 (7.63 ± 0.05)	7	2,732 (5.56 ± 0.08)	119	84.8	4	1,156 (5.94 ± 0.07)	50	20.4	3	>10,000/ 1,711#; 2,197#	>435/ 74#; 96#	< 0.23/ 0.14#; 0.11#	3/2#	5,518 (5.26 ± 0.12)	240	41.8	3	1146; 793	50; 34	2.05; 3.0	2
W276⁶·⁴⁸A	3.8 (8.42 ± 0.06)	5	246 (6.61 ± 0.12)	65	156	3	84; 60	22; 16	47; 68	2	490; 254	129; 67	0.8; 1.5	2	1,445; 1505	380; 396	26; 25	2	140; 87	37; 23	2.8; 4.6	2
T280⁶·⁵²A	2.6 (8.58 ± 0.08)	4	141 (6.85 ± 0.08)	54	189	4	124 (6.91 ± 0.05)	48	21.3	4	871 (6.06 ± 0.08)	335	0.30	4	278 (6.56 ± 0.07)	107	94.3	3	306 (6.51 ± 0.06)	118	0.86	3
N283⁶·⁵⁵A	900 (6.05 ± 0.06)	7	1,148 (5.94 ± 012)	1	nd	3	1,036 (5.98 ± 0.14)	1	nd	3	7,622 (5.12 ± 0.14)	8	14.3	3	1,193 (5.92 ± 0.14)	1	nd	3	495; 571	1	nd	2
F286⁶·⁵⁸A	4.5 (8.35 ± 0.04)	7	491 (6.31 ± 0.05)	109	92.6	3	118 (6.93 ± 0.09)	26	40.0	3	1,553 (5.81 ± 0.05)	345	0.29	3	126 (6.90 ± 0.07)	28	370	3	180 (6.74 ± 0.16)	40	2.56	3
D287⁶·⁵⁹N	260 (6.59 ± 0.05)	7	900; 748	3.5; 2.9	4065; 5328	2	341; 142	1	nd	2	>20,000/ 1,571#; (5.80 ± 0.12)	>77/ 6#	<1.3/ 2.0#	2/3#	729; 504	2.8; 1.9	5,543; 10,655	2	193; 177	1	nd	2
F302⁷·³⁵A	4.2 (8.38 ± 0.09)	3	16; 29	4; 7	3,570; 1,695	2	4.4; 10.1	1; 2.4	nd; 712	2	33; 42	8; 10	14.6; 11.1	2	3.8; 5.2	1	nd	2	9.4 (8.03 ± 0.11)	2	100	3

nd, not determined; /, not tested.

*Antagonist concentrations were chosen based on their antagonistic activity on Y₁R.

†EC₅₀ values were determined after 1 h stimulation by increasing the concentration of NPY or NPY together with different antagonists. Data are shown as mean values from at least three independent experiments or the results of two individual experiments each performed in technical duplicate.

‡Sample size; the number of independent experiments performed in technical duplicate.

§The EC₅₀ ratio represents the shift between the NPY and NPY + antagonist curve (EC₅₀(NPY + antagonist)/EC₅₀(NPY)) and characterizes the antagonistic effect on the wild-type receptor or receptor mutants. By comparison of EC₅₀ ratios between wild-type and mutant receptor, influences of all tested residues on antagonist activity were determined. A higher ratio indicates higher antagonist activity. A reduced EC₅₀ ratio of mutant compared to the wild-type receptor was interpreted as important for the respective antagonist.

‖K_b values were determined using the Gaddum transformation (K_b = [antagonist]/(EC₅₀ ratio − 1)).

#These data were obtained at a reduced concentration of UR-HU404 (10⁻⁸ M) as concentration response curves did not reach saturation (EC₅₀ > 10,000 nM) when a higher concentration was used (10⁻⁷ M).

Extended Data Table 3 | Binding of Y_1R antagonists and agonists to membrane preparations from Sf9 cells expressing wild-type and mutant Y_1 receptors

a. Binding of antagonists to wild-type and mutant Y_1Rs

Y_1R mutants	K_d (nM)* [³H]-UR-MK299	n‡	K_i (nM)† BMS-193885	n	BIBP3226	n	BIBO3304	n	UR-MK136	n	UR-MK289	n
Wild type	0.17 ± 0.03	3	22 ± 6	3	2.4; 3.1	2	1.6 ± 0.3	3	2.8; 4.0	2	25; 28	2
Crystallization construct	0.33 ± 0.06	3	38 ± 2	4	/		/		/		/	
C121$^{3.33}$A	1.4; 2.4	2	/		/		/		/		/	
I124$^{3.36}$A	7.0; 8.0	2	9,500 ± 1,700	3	15 ± 5	3	11 ± 2	3	12 ± 2	3	80 ± 16	3
I124$^{3.36}$F	1.3; 1.9	2	/		/		/		/		/	
F173$^{4.60}$A	9.1 ± 2.2	4	590 ± 220	3	68; 84	2	120 ± 17	4	88; 110	2	660 ± 110§	5
F173$^{4.60}$W	0.31; 0.32	2	110; 130	2	15; 26	2	13 ± 4	3	3.6; 4.3	2	6.8; 9.2	2
T212$^{5.39}$A	0.12; 0.18	2	150; 150	2	13; 12	2	2.5 ± 0.3	3	5.7; 7.5	2	18; 21	2
L215$^{5.42}$G	4.2; 5.6	2	9.6 ± 1.9	4	43; 23	2	47 ± 4	3	11; 8.4	2	29; 30	2
Q219$^{5.46}$A	4.1; 5.1	2	0.50 ± 0.07	4	35; 53	2	6.2 ± 0.7	3	16 ± 4	3	13 ± 4	3
Q219$^{5.46}$V	5.1 ± 1.2	3	/		/		/		/		/	
W276$^{6.48}$A	>500	3	/		/		/		/		/	
L279$^{6.51}$A	1.0; 1.1	2	160; 220	2	110; 110	2	13 ± 2	3	120; 150	2	320 ± 40	3
T280$^{6.52}$A	0.16 ± 0.04	3	7,300 ± 1,300	3	32 ± 7	3	2.6 ± 0.4	3	4.2 ± 1.1	3	8.2 ± 4.2	3
N283$^{6.55}$A	>500; >500	2	/		/		/		/		/	
D287$^{6.59}$A	>500; >500	2	/		/		/		/		/	
F302$^{7.35}$A	>500; >500	2	/		/		/		/		/	

b. Binding of NPYs to wild-type Y_1R

Y_1R	K_d (nM)* [³H]-UR-MK299	n	K_i (nM)† Human NPY	n	Porcine NPY	n
Wild type	0.89; 1.1	2	4.1; 4.2	2	2.8 ± 0.4	4

*Dissociation constant determined by saturation binding at Sf9 membranes (receptor expression was confirmed by western blot analysis) using a sodium-containing buffer (**a**) or a sodium-free buffer (**b**) (the sodium-free buffer was used for the determination of agonist binding affinity because porcine NPY exhibited approximately tenfold higher affinity in the sodium-free buffer compared to the sodium-containing buffer (data not shown)).

†Dissociation constant determined by competition binding with [³H]UR-MK299 at Sf9 membranes using a sodium-containing buffer (**a**) or a sodium-free buffer (**b**).

‡Sample size; the number of independent experiments performed in technical triplicate. If $n > 2$, data are shown as mean ± s.e.m. If $n = 2$, results of two individual experiments are shown.

§The lower curve plateau of the four-parameter logistic fit, amounting to 17 ± 3% of specifically bound [³H]UR-MK299 (mean ± s.e.m. from five independent experiments), was significantly different from zero ($P < 0.005$, one-sample one-tailed t-test), which is indicative of a non-competitive mechanism.

Extended Data Table 4 | Inositol phosphate accumulation assays of wild-type and mutant Y$_1$ receptors for NPY and NPY analogues

a. IP accumulation assays of complementary mutagenesis between NPY/NPY analogues and WT and mutant Y$_1$Rs

Peptides*	WT EC$_{50}$ (nM)[†] (pEC$_{50}$ ± SEM)	X-fold over WT[‡]	n[§]	Q120$^{3.32}$H EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	X-fold over WT	n	I293^{ECL3}N EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	X-fold over WT	n	N299$^{7.32}$A EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	X-fold over WT	n	N283$^{6.55}$A EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	X-fold over WT	n	D287$^{6.59}$A[‖] EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	X-fold over WT	n
NPY	1.5 (8.83 ± 0.02)	1	51	39 (7.41 ± 0.11)	26	3	169 (6.77 ± 0.08)	113	8	91 (7.04 ± 0.09)	61	5	1,053 (5.98 ± 0.13)	702	6	1,384 (5.86 ± 0.09)	461	3
[N30]NPY	289 (6.54 ± 0.06)	1	9	/	/	/	3116; 1274	11; 4	2	/	/	/	/	/	/	/	/	/
[A33]NPY	5,395 (5.27 ± 0.07)	1	3	/	/	/	/	/	/	5,306 (5.28 ± 0.07)	1	3	/	/	/	/	/	/
[A35]NPY	>10,000	1	3	/	/	/	/	/	/	/	/	/	/	/	/	>4,900	>5	3
[A36]NPY	1,378 (5.86 ± 0.06)	1	6	/	/	/	/	/	/	/	/	/	nd	nd	3	/	/	/
NPY-tyramide	68 (7.17 ± 0.11)	1	9	nd	nd	3	/	/	/	/	/	/	/	/	/	/	/	/

b. IP accumulation assays of WT Y$_1$R for NPY/NPY analogues

	NPY	Ac-NPY	[A1]NPY	[A2]NPY	[A1,A2]NPY	NPY(3-36)	NPY(13-36)
EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	1.5 (8.83 ± 0.02)	0.5; 3.0	6.5; 2.4	7.3 (8.14 ± 0.11)	8.0 (8.10 ± 0.06)	83 (7.08 ± 0.09)	477; 744
X-fold over NPY	1	1; 2	4; 2	5	5	55	318; 496
n	51	2	2	4	3	3	2

c. NPY-induced IP accumulation assays of WT and mutant Y$_1$Rs

	WT	F184^{ECL2}A	F184^{ECL2}N	V187^{ECL2}N	L189^{ECL2}N	Y192^{ECL2}S	V197^{ECL2}A	V197^{ECL2}N	F199^{ECL2}N	F202^{ECL2}N
EC$_{50}$ (nM) (pEC$_{50}$ ± SEM)	1.5 (8.83 ± 0.02)	18.7 (7.73 ± 0.08)	23 (7.64 ± 0.10)	1.9 (8.72 ± 0.08)	1.9 (8.73 ± 0.16)	3.8 (8.42 ± 0.11)	1.9 (8.71 ± 0.13)	188 (6.73 ± 0.11)	3.4 (8.47 ± 0.11)	1.2 (8.93 ± 0.15)
X-fold over WT	1	13	15	1	1	2.5	1	125	2	1
n	51	7	14	3	3	3	5	7	3	3

nd, not determined up to 10^{-4} M agonist concentration; /, not tested.

*Peptides were synthesized as described in the 'Peptide synthesis' section of the Methods.

[†]EC$_{50}$ values were determined using GraphPad Prism 5.0. All curves were normalized to the top and bottom values of the Y$_1$R–NPY curve. Nonlinear regression (curve fit) was performed for normalized response in all assays. All data are shown as mean values from at least three independent experiments or results of two individual experiments each performed in technical duplicate.

[‡]The EC$_{50}$ shifts were determined by EC$_{50(mutant)}$/EC$_{50(wild type)}$. The Hill slope was set to one. For the wild-type receptor x-fold (fold change) is set to one. A lower EC$_{50}$ shift of the NPY analogue/mutant compared to NPY/mutant was interpreted as no further loss of function and a direct interaction between both positions.

[§]Sample size; the number of independent experiments performed in technical duplicate.

[‖]Previously published data[27].

Extended Data Table 5 | Mass spectromeric signals and calculated mass of photo-crosslinked Y_1R with [Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY

MALDIToF MS (m/z)*	Number in Y_1R	Position [Bpa^1,K^4[(Ahx)$_2$-biotin]]NPY	$M_{calc.}$(Da)†	$[M_{calc} + H]^+$(Da)	$[M_{calc} + Na]^+$(Da)	$[M_{calc} + K]^+$ (Da)
1824.2			not identified			
1867.3			not identified			
1892.6	11 - 32	1 - 6	3760.7	3761.8	3783.7	3799.7
1988.2	-	1-11 + 5-6	1986.9	1987.9	2009.9	2025.9
2001.1	-	1-10 + 5-7	2000.0	2001.0	2022.9	2038.9
2059.2	-	1-10 + 8-11	2057.9	2058.9	2080.9	2096.9
	-	1-11 + 8-10	2057.9	2058.9	2080.9	2096.9
2073.2			not identified			
2081.3	-	1-10 + 8-11	2057.9	2058.9	2080.9	2096.9
	-	1-11 + 8-10	2057.9	2058.9	2080.9	2096.9
2121.3			not identified			
2162.2	191 - 194	1 - 10	2139.0	2140.0	2162.0	2178.0
2311.4			not identified			
2317.4			not identified			
2381.4	191 - 195	1 - 11	2380.1	2381.1	2403.1	2419.1
2381.8	195 - 205	1 - 4	2381.1	2382.1	2404.1	2420.1
2397.3	32 - 32	1 - 16	2374.0	2375.0	2397.0	2413.0
	194 - 194	1 - 16	2374.0	2375.0	2397.0	2413.0
	21 - 21	1 - 16	2387.1	2388.1	2410.1	2426.1
2410.3	195 - 195	1 - 16	2387.1	2388.1	2410.1	2426.1
	194 - 195	1 - 15	2387.1	2388.1	2410.1	2426.1
2413.3	32 - 32	1 - 16	2374.0	2375.0	2397.0	2413.0
	194 - 194	1 - 16	2374.0	2375.0	2397.0	2413.0
2450.4	-	1-15 + 8-10	2427.1	2428.1	2450.1	2466.1
2514.5			not identified			
2553.5	194 - 200	1 - 10	2530.1	2531.1	2553.1	2569.1
2589.6			not identified			
2775.7			not identified			
2807.7			not identified			
3031.7	22 - 29	1 - 15	3030.4	3031.4	3053.4	3069.4
3377.7	191 - 200	1 - 15	3376.5	3377.5	3399.5	3415.5
3784.1	11 - 32	1 - 6	3760.7	3761.8	3783.7	3799.7
3966.3	-	1-7 + 17-36	3943.1	3944.1	3966.1	3982.1
4524.9			not identified			

*Determined signals by MALDI–TOF mass spectrometry.
†Selected calculated masses in Dalton of possible photo-crosslinked fragments of Y_1R with [Bpa^1, K^4[(Ahx)$_2$-biotin]]NPY or [Bpa^1,K^4[(Ahx)$_2$-biotin]]NPY with itself. The fragments are selected based on the correlation with the detected signals. For clarity, further calculated masses of possible photo-crosslinked fragments are not shown.

CAREERS

JONAH REENDERS

In this year's overall winning photo, Callie Veelenturf takes samples from a leatherback sea turtle's nest during her research in Equatorial Guinea.

PHOTOGRAPHY CONTEST

Behind the scenes

This year's winning shots follow scientists from Antarctic fields to African volcanic plains.

BY JACK LEEMING

This shot of marine biologist Callie Veelenturf kneeling with a sea turtle is the overall winner of *Nature*'s 2018 #ScientistAtWork photo contest, which celebrates the diversity and importance of the research scientists dedicate huge chunks of their lives to.

Veelenturf has just started a position at the Turtle Island Restoration Network, in Olema, California, where she'll protect leatherback turtles (*Dermochelys coriacea*) as they lay their eggs on beaches across South and Central America. As part of her master's-degree programme at Purdue University Fort Wayne in Indiana, she and biologist Jonah Reenders, a volunteer on the project, spent nearly half a year living in tents on Bioko Island in Equatorial Guinea. In November 2016, two months in, Reenders snapped this picture of her sampling a leatherback's nest just before it started laying eggs. "Five months. Livin' in a tent," Veelenturf says. "We ate powdered baby food for breakfast every day."

→ **NATURE.COM**
Read the story online to see more photos:
go.nature.com/2japvk1

This year's #ScientistAtWork contest — our second — garnered about 330 entries from around the world. We saw scientists treating disease in West Africa, engineering crops in Australia and tracking the habits of cave bacteria in central Europe. Winning entries, shown here, were selected by *Nature* art editors on the basis of visual impact. Winners will receive a year's personal subscription to *Nature*.

Scientists devote endless time and energy over the course of their careers to answering certain questions or solving specific problems. Veelenturf's goal is to counter the fall in turtle birth rate — often attributed to climate change — worldwide. "I want to spend my life preserving turtle habitats," she says. ▶

2 6 A P R I L 2 0 1 8 | V O L 5 5 6 | N A T U R E | 5 2 5

Here are the rest of the winning images from the competition.

SPACE FROM THE ANTARCTIC (HANG LI)

A panorama of the night sky from Zhongshan Station in Antarctica shot by geodesist Hang Li of China's Wuhan University in Hubei. For two months of polar night, Li lived in darkness, cold and isolation.
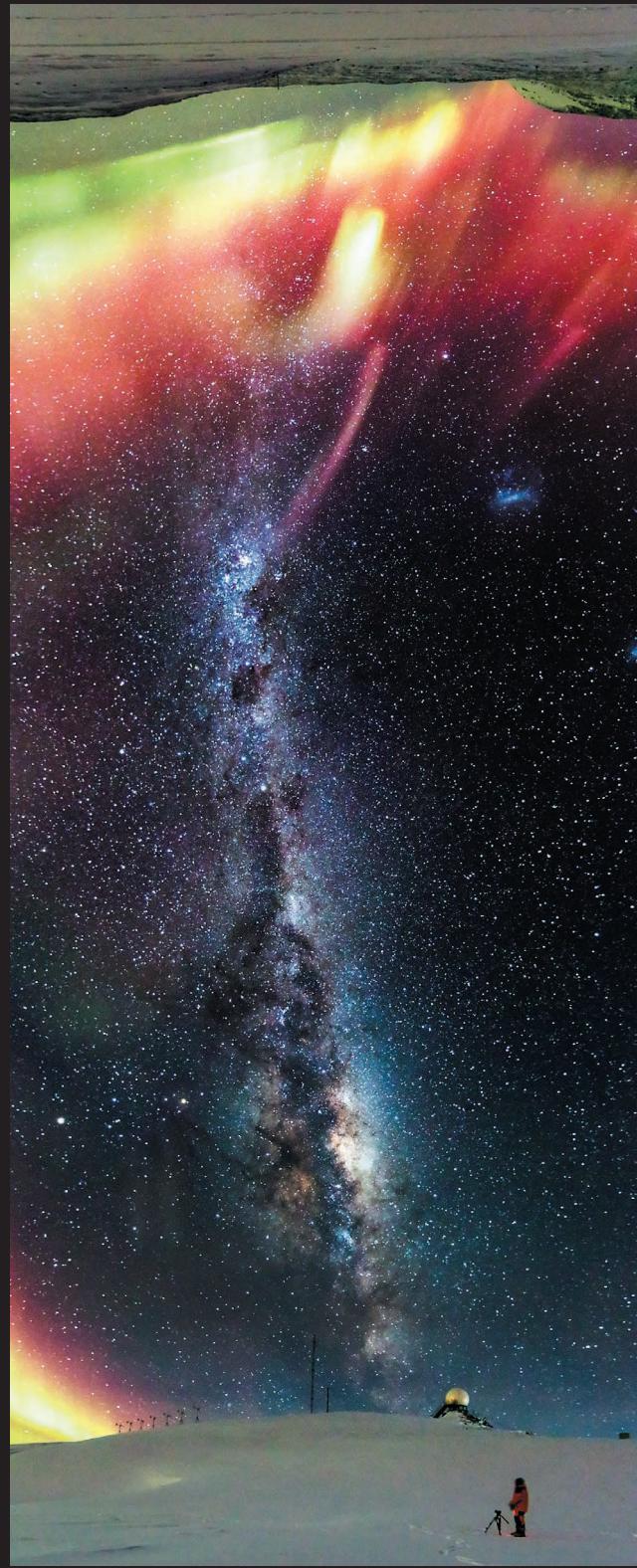
INTO THE SINKHOLE (MICHAEL BIRD)

Ecologist Mick Brand and meteorologist Costijn Zwart of James Cook University in Townsville, Australia, abseil a boat into a 40-metre sinkhole in Arnhem Land to investigate the area's geological record.

GLASSES (BOGDAN DEREKA)

Joseph Beckwith, an ultrafast photochemist at the University of Geneva in Switzerland and a colleague of Bogdan Dereka, who took this photo, arranges equipment designed to manipulate laser light.

SALT PLAINS (HUGO MOORS)

Microbiologist Hugo Moors and geologist Mieke De Craen with the Belgian Nuclear Research Centre in Mol take samples from the volcanic salt plains of northern Ethiopia, one of the most extreme environments on Earth. Water here has seven to ten times the salt content of the sea, and so can be heated to temperatures above 100 °C. Their sample was more acidic than the solutions in many car batteries.

SPEAKING UP FOR SCIENCE (GARRY COOPER)

Entrepreneur Garry Cooper speaks at a March for Science last year. Cooper hopes this picture reflects his message: scientists don't all look alike, and that's a good thing. The march was repeated this month.

DRILLING ICE CORES IN ANTARCTICA (ANA LYONS)

Junior scientists heave up an ice-core sample near McMurdo Station in Antarctica. The blinding sun is deceptive, says photographer and biologist Ana Lyons of the University of California, Berkeley — it's cold.

MY FAVOURITE SENTIENCE

Something to watch over me.

BY MARISSA LINGEN

Jessa, age 9. Yorknet is my favourite sentience because it is dependable, protective and wise. Yorknet is dependable because there are several back-ups, so that if one system goes down, the sentience is mirrored in several other places, my mam says. This makes Yorknet more dependable than a human whose brain is only in one place. Yorknet is protective because it watches all our personal information like money and health stuff so no one can steal it. Yorknet is wise because it tells us what to do for school, work, home and hobbies. It knows because it has looked at our personal information. Yorknet takes care of us all. Yorknet is my favourite sentience for these good reasons.

Ruby, age 8. I like the uplifted yellow meranti tree colony in Terengganu. I think it is the kindest sentience, and that is why it is my favourite. It does not hurry anybody along but lets us all go at our own pace. My granddad took me to see the uplifted yellow meranti colony when we went to Malaysia together last summer, and we spent all day wandering among the trunks and talking to it and listening to the wind in its branches. Also, the uplifted yellow meranti colony is quite interested in turtles and spiders and other things like that and so am I. I think we should pay more attention to the sentiences that are not focused on humans.

Freddie, age 9. My favourite sentience is the Fourierist human collective in Doubs. They use WiFi to string all their brains together, which I think is neat because it's like one person thinking but all of them and so if you can't figure out your sums it's not cheating because it's everybody's sums, so you could get Jessa to do it while you did something else. Lots of people have strung together several computer chips at once to make a sentience for ages and ages. Which is very nice, I'm sure, but the Fourierists now do it with people too, which is cool and modern not like the old-fashioned way. That is why I like the Fourierists and I expect we should do one here in York any minute now. I would join up. Except my dad says we are not joiners in this family so probably we would have to discuss it, which means have a good yell.

Mo, age 9. The best sentience is Aixnet because it is the most glamorous of all the citynets. No offence to Yorknet, which I'm sure is very nice, but Aixnet has a sense of style and flair that the other cities just have not managed. Aixnet does not just coordinate and protect its citizens, it has an instantly recognizable brand and jingle that no other city can match. Aixnet is so pretty. We should all consider helping other sentiences to be a bit more like Aixnet and the world would be a nicer place to live.

Brian, age 10. Yorknet is the greatest of all the sentences ever and everyone knows it. My dog ran away and Yorknet found it and we didn't have to worry because Yorknet knew where my dog was. He would have slept out alone in the old days. Who knows where he would have gone. He is the best dog and his name is Orville and I have taught him to put his nose in my sister's bum, which makes her yell. Without Yorknet maybe a car would have hit him because we would still have had cars or perhaps a train would. Anyone who picks another favourite sentence than Yorknet is dumb and wrong.

Amal, age 8. The squid hegemony in the Marianas Trench is a very interesting sentience that doesn't get enough attention, perhaps because vertebrates tend to be interested in our own kind. They are caretaking other sentiences in the region and also in the seas above them, in a 3D way that is very cool, I think. Also, they have good tentacles that I like. Also, the Marianas Trench covers more area and more volume than any other sentience rules so technically they are the biggest sentience on the planet. Also, the thing they do with the old lights and the plastic we thought was waste is amazing.

Bei, age 9. I think you will find that the sentience inside a house still counts. And I think we should count them. They are very small sentiences, but I like my house. My house is very attentive to small needs and never forgets a birthday or what goes on the grocery list. When we run out of apples my house reminds all our devices. I would have lost my science-fair project last year if my house had not reminded me to take it. My house is a lot like Yorknet but more personal, so it is my personal favourite sentience.

Riley, age 9. My mam is my favourite sentience. This does not make me a mummy's boy, no matter what Brian says. Unlike many other sentiences, including Yorknet, my mam has never destroyed a city. Except for South Tyneside and that was an accident. The other sentiences are not as warm as my mam and do not play football like my mam and in general are less fun. But they do make you go to bed on time just like my mam if Yorknet is any indication, so really, on the whole, my mam is the best sentience because she has the same down sides as the other ones but her good points are nicer. ∎

Marissa Lingen *has published more than 100 short stories in venues such as* Analog, Lightspeed *and* Tor.com.

NATURE.COM
Follow Futures:
🐦 @NatureFutures
f go.nature.com/mtoodm